

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

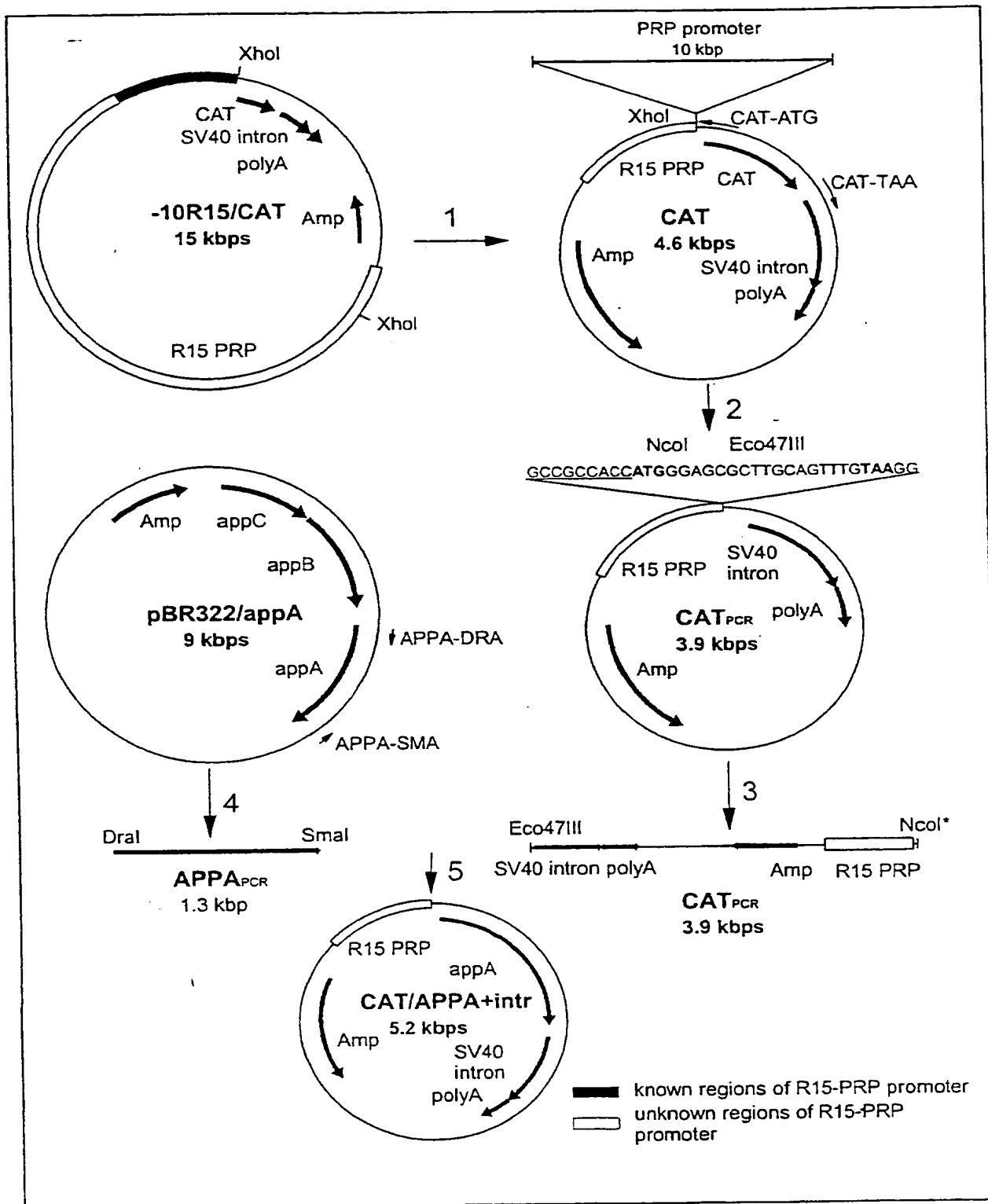
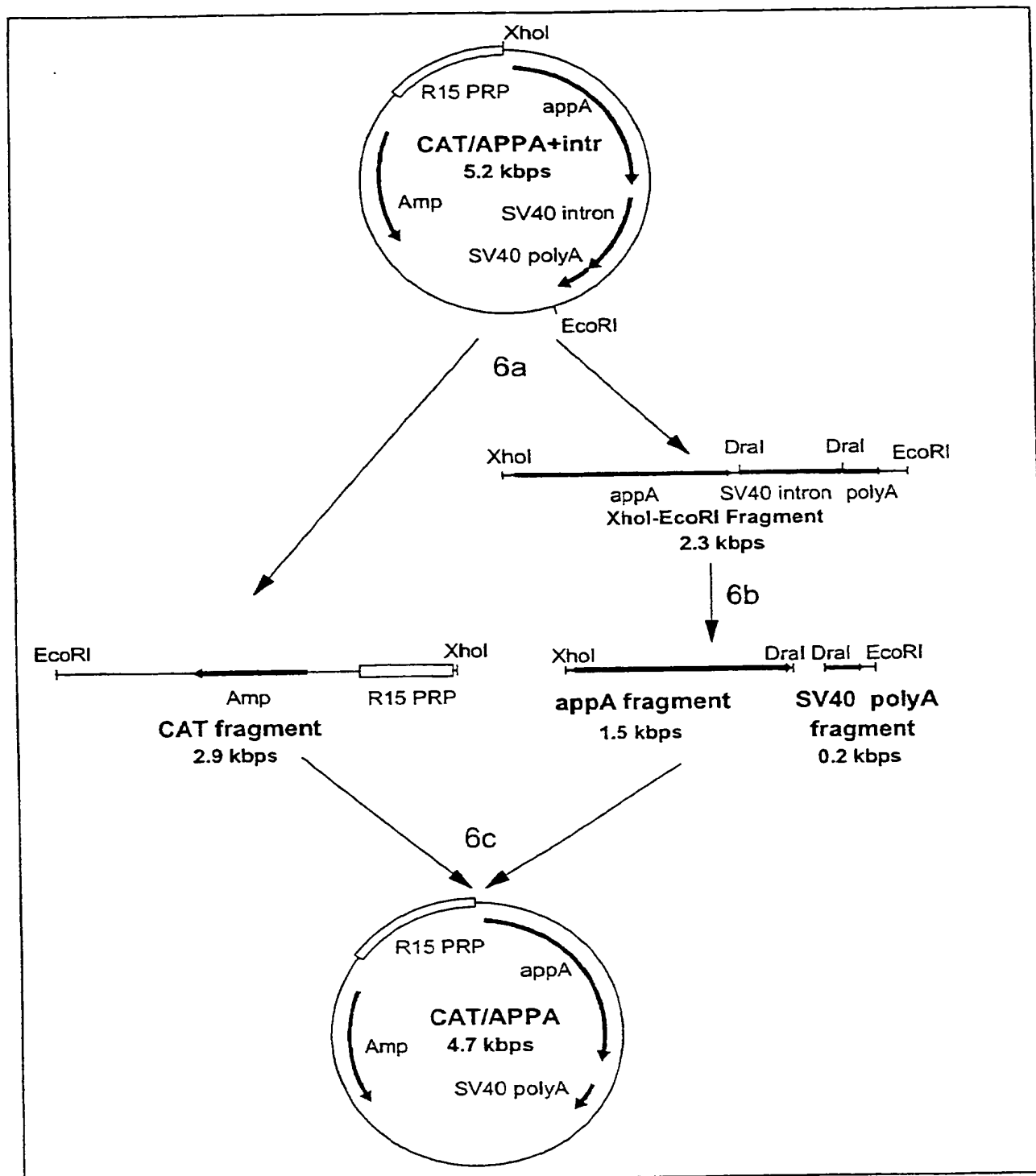


Figure 1

**Figure 1 (continued)**

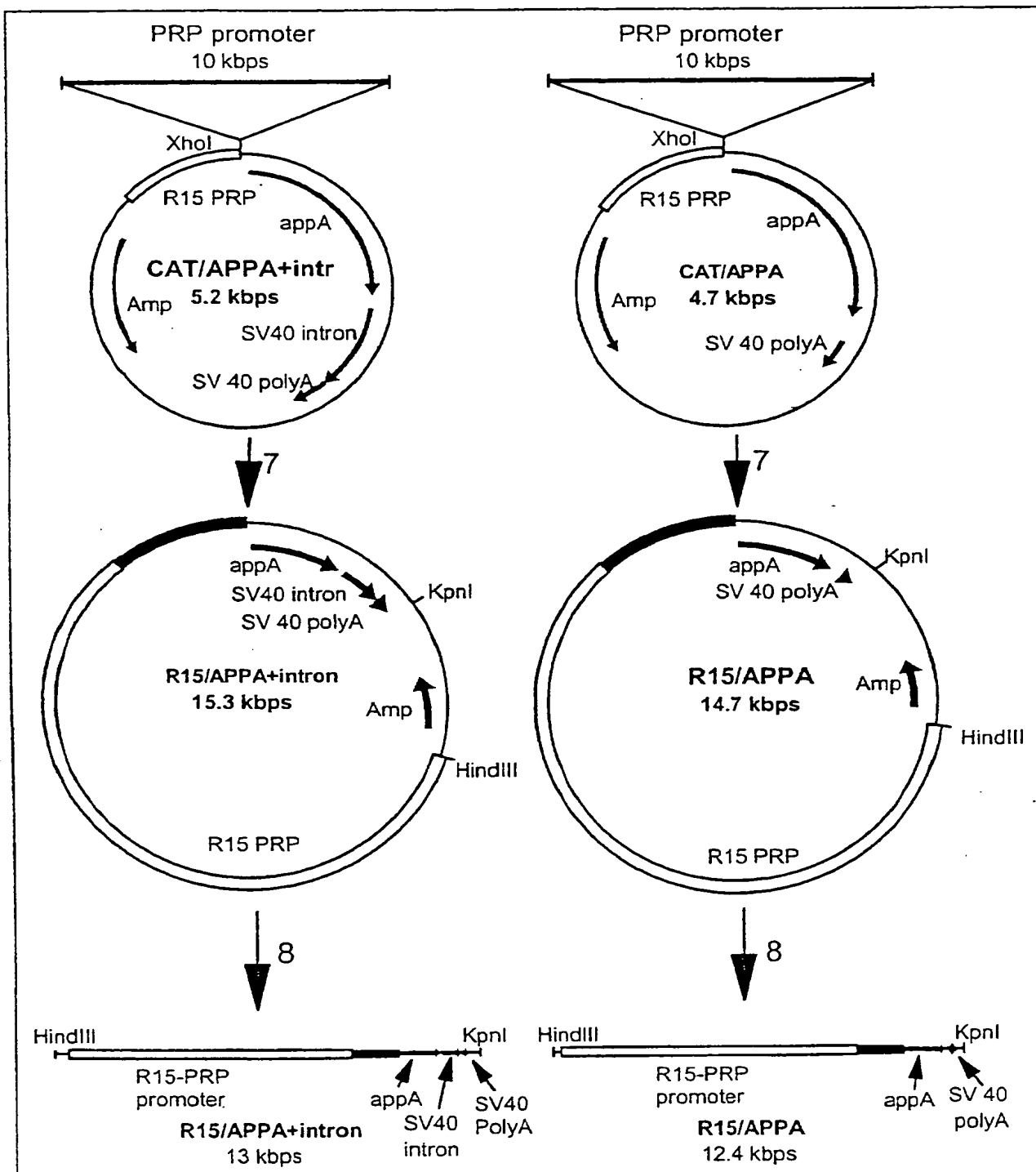
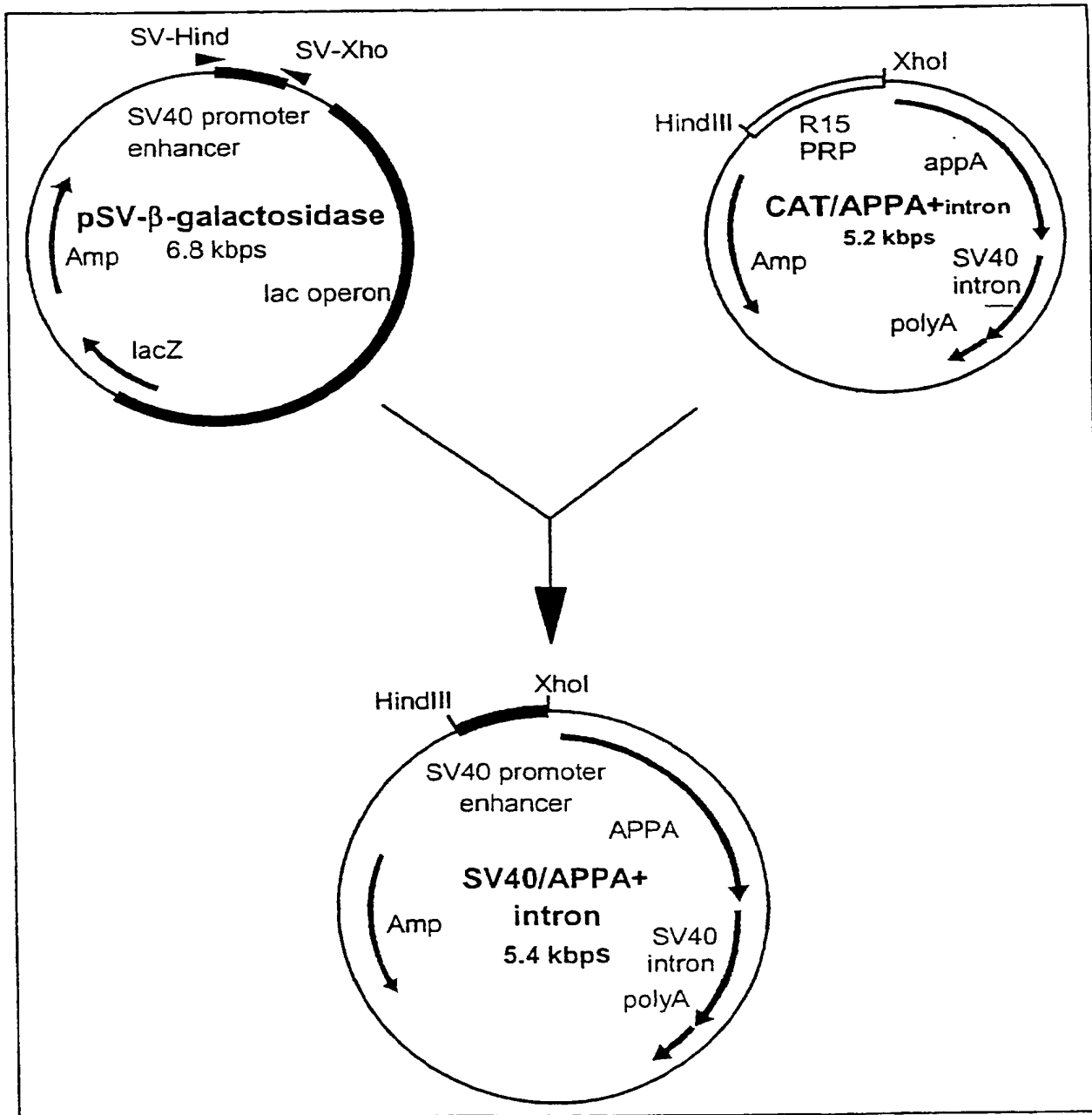
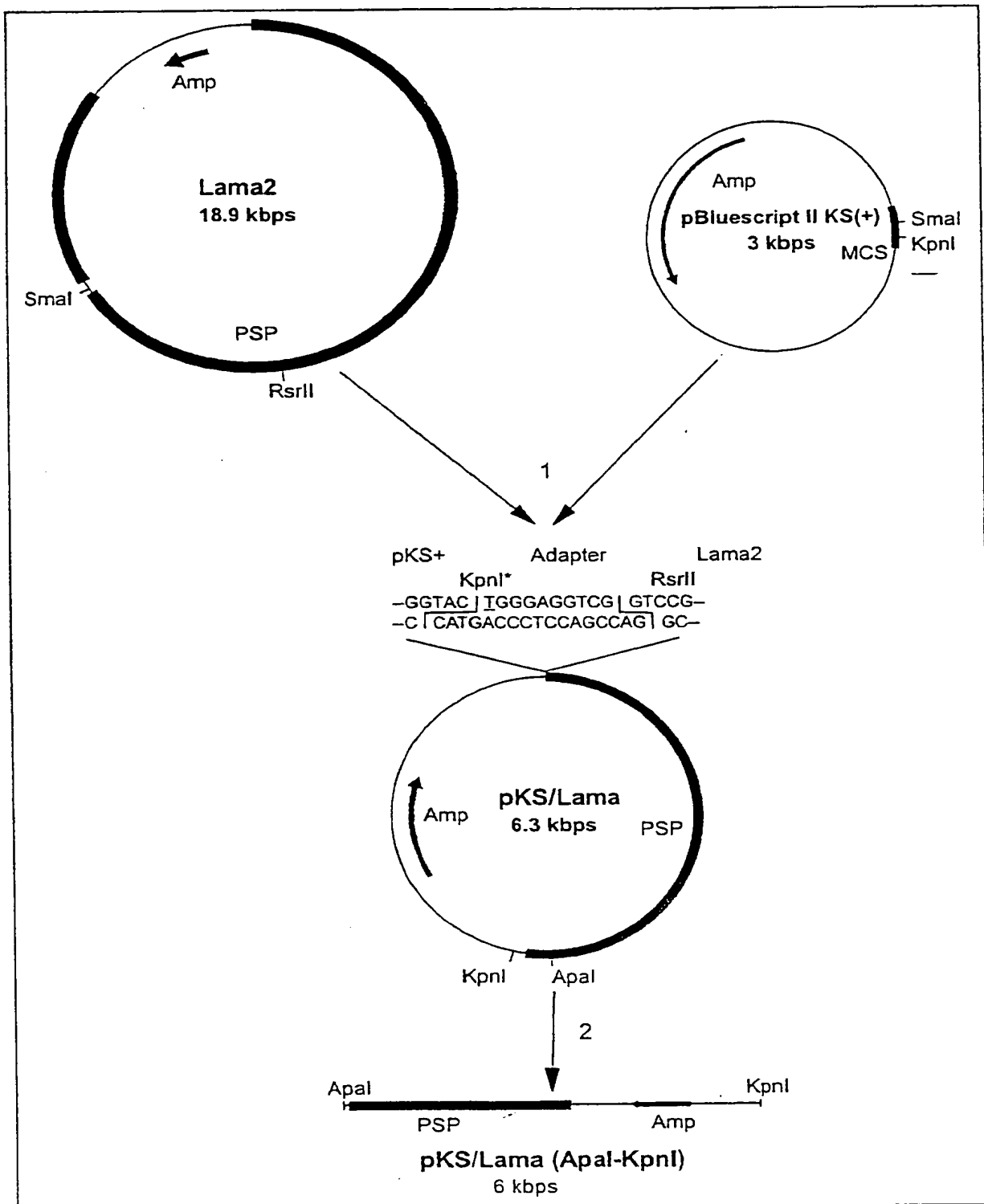


Figure 1 (continued)

**Figure 2**

**Figure 3**

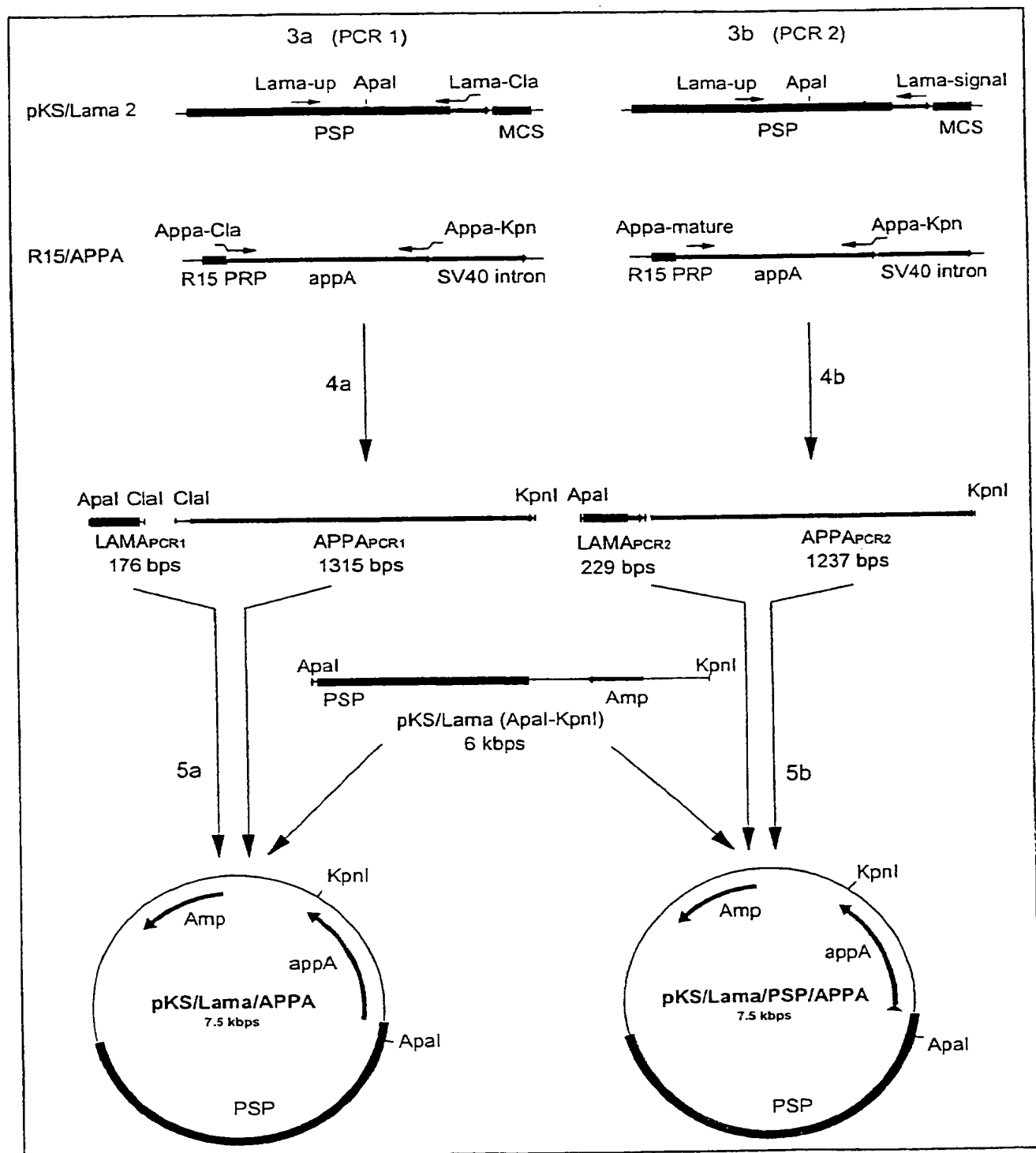


Figure 3 (continued)

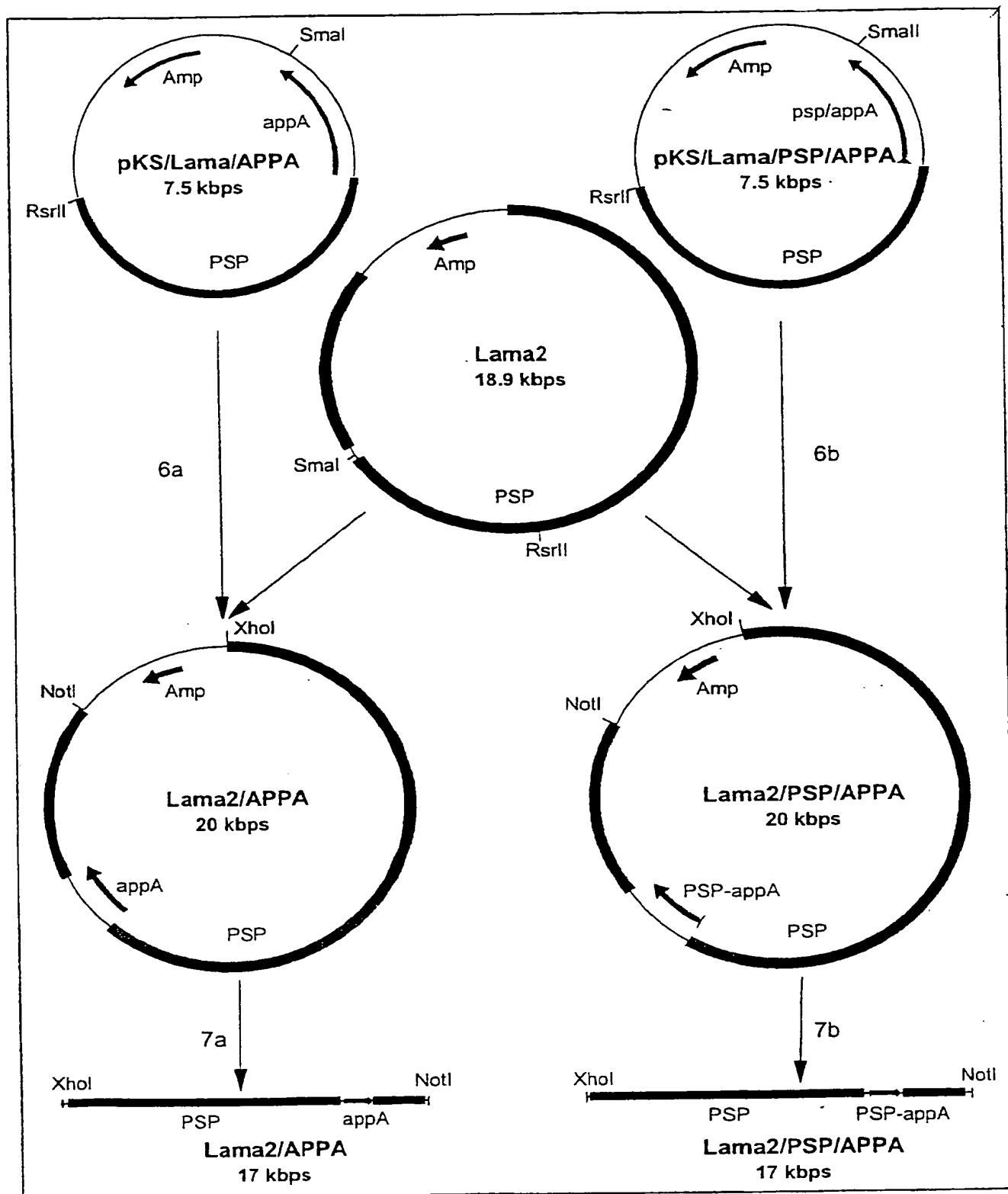


Figure 3 (continued)

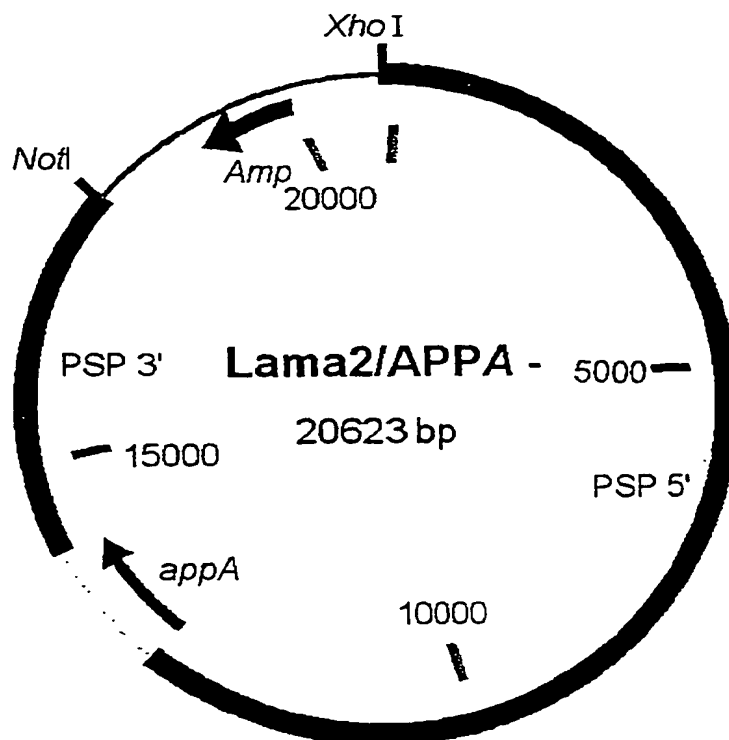


Figure 4. Schematic diagram of the Lama2/APP A construct.

Figure 5. The nucleic acid sequence of the Lama2/APPA plasmid (SEQ ID NO: 1)

LOCUS Lama-appA 20623 bp DNA CIRCULAR SYN 17-JAN-2000
 DEFINITION Lama 2/APPA transgenic construct
 ACCESSION Lama 2-appA,
 KEYWORDS parotid secretory protein; acid glucose-1-phosphatase; appA
 gene;
 periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 cloning vector
 REFERENCE 1 (bases 1 to 20623)
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.
 FEATURES
 DEFINITION M. musculus Psp gene for parotid secretory protein.
 ACCESSION X68699
 VERSION X68699.1 GI:53809
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 3777 to 5332;)
 AUTHORS Svendsen, P., Laursen, J., Krogh-Pedersen, H. and Hjorth, J.P.
 TITLE Novel salivary gland specific binding elements located in the PSP
 proximal enhancer core
 JOURNAL Nucleic Acids Res. 26 (11), 2761-2770 (1998)
 MEDLINE 98256451
 REFERENCE 2 (bases 7147 to 12653; 13952 to 17731)
 AUTHORS Mikkelsen, T.R.
 TITLE Direct Submission
 JOURNAL Submitted (07-OCT-1992) T.R. Mikkelsen, Department of Molecular
 Biology, University of Aarhus, CF Mollers Alle 130, 8000
 Aarhus, DENMARK
 REFERENCE 3 (bases 7147 to 12653; 13952 to 17731)
 AUTHORS Laursen J, Hjorth JP
 TITLE A cassette for high-level expression in the mouse salivary glands.
 JOURNAL Gene 1997 Oct 1;198(1-2):367-72
 MEDLINE 9370303

FEATURES Location/Qualifiers
 source 1.to 12653; 13952 to 17731
 /organism="Mus musculus"
 /strain="C3H/As"
 /db_xref="taxon:10090"
 /chromosome="2"
 /map="Estimate: 69 cM from centromere"
 /clone="Lambda YP1, Lambda YP3, Lambda YP7"
 /clone_lib="Lambda-PHAGE (Lambda L47.1)"
 /germline
 /note="Allele: b"
 misc_feature 3777-5332
 /gene="PSP"
 /function="salivary gland specific positive acting
 regulatory region"
 enhancer 7147..8724
 /evidence=experimental
 exon 11778..11824
 /gene="Psp"
 /note="exon a"
 /number=1
 /evidence=experimental
 exon 12626.. 14190
 /gene="Psp"
 /note="exon b fused with exons h and i"
 misc_feature 12644-12652

Figure 5 (continued):

```

                /function=" consensus sequence for initiation in higher
                eukaryotes "
misc_feature 13952-13965
                /function=" M13mp18 polylinker"

DEFINITION  E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene,

ACCESSION   M58708 L03370 L03371 L03372 L03373 L03374 L03375
VERSION     M58708.1 GI:145283
SOURCE      Escherichia coli DNA.
ORGANISM    Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.

REFERENCE   1 (bases 12653..13951)
AUTHORS    Dassa,J., Marck,C. and Boquet,P.L.
TITLE      The complete nucleotide sequence of the Escherichia coli gene appA
            reveals significant homology between pH 2.5 acid phosphatase
            and glucose-1-phosphatase
JOURNAL    J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE    90368616

FEATURES             Location/Qualifiers
     Source                12653..13951
                            /organism="Escherichia coli"
                            /db_xref="taxon:562"
     sig_peptide           12653..12718
     /gene="appA"
     CDS12653              13951
                            /gene="appA"
                            /standard_name="acid phosphatase/phytase"
                            /transl_table=11
                            /product="periplasmic phosphoanhydride phosphohydrolase"
                            /protein_id="AAA72086.1"
                            /db_xref="GI:145285"

     /translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
TKATQLMQDVTPDAWPTWPVKLGWLTTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS
GOVAIIADVDERTKRTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHOWNTLLSLHNAQF
YLLQRTPEVARSRATPLLDLIKALTTPHPPQKQAYGVTLPSTVLFIAGHDTNLANLGG
ALELNWTLPGQPDNTTPPGGELVFERWRRLSDNSQWIIQVSLVFQTLQQMRDKTPLSLNT
PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"
     mat_peptide           12719 13948
                            /gene="appA"
                            /product="periplasmic phosphoanhydride phosphohydrolase"

     mutation              replace(12659.. 12661,"gcg changed to gcc")
                            /gene="appA"
                            /standard_name="A3 mutant"
                            /note="created by site directed mutagenesis"
                            /citation={3}
                            /phenotype="silent mutation"
     mutation              replace(13934..13936," ccg changed to ccc")
                            /gene="appA"
                            /standard_name=" P428 mutant"
                            /note="created by site directed mutagenesis"
                            /citation={3}
                            /phenotype=" silent mutation "
     mutation              replace(13937..13939," gcg changed to gct")
                            /gene="appA"
                            /standard_name=" A429 mutant"
                            /note="created by site directed mutagenesis"

```

Figure 5 (continued):

```

/citation=[3]
/phenotype=" silent mutation "

DEFINITION      pBluescript II KS(+) vector DNA,
ACCESSION        X52327
VERSION          X52327.1  GI:58061
KEYWORDS         artificial sequence; cloning vector; expression vector; vector.
SOURCE           synthetic construct.
ORGANISM         synthetic construct
                 artificial sequence.

REFERENCE        1      (bases 17732 to 20623)
AUTHORS          Thomas,E.A.
TITLE            Direct Submission
JOURNAL          Submitted (20-FEB-1990) Thomas E.A., Stratagene Cloning
                 Systems, 11099 North Torrey Pines Rd., La Jolla, CA 92037, USA

REFERENCE        2      (bases 17732 to 20623)
AUTHORS          Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
TITLE            Lambda ZAP: a bacteriophage lambda expression vector with in
                 vivo excision properties
JOURNAL          Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE          88319944
REFERENCE        3      (bases 17732 to 20623)
AUTHORS          Altling-Mees,M.A. and Short,J.M.
TITLE            pBluescript II: gene mapping vectors
JOURNAL          Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE          90067967
FEATURES         Location/Qualifiers
Source           17732 to 20623
                 /organism="synthetic construct"
                 /db_xref="taxon:32630"
CDS              complement (18967..19827)
                 /gene="Amp"
                 /product="b-lactamase"

```

BASE COUNT 5449 a 4847 c 4902 g 5424 t

ORIGIN

```

1  TCGAGAGTAT CTTTGTCAAG TGTGCCTCCA ACAAAGGGGT ACTGTTGCCC ACATAGAAAAG
61 ATCTAAACTA ATTAATTAAT CCCTCACCCG CAAATCTTTC AGTCACTAAG TTAGCACGAT
121 TGTGTAACAA GTTCTCCAAA GGAGAGATAC AGATGAGTGC GTATAGGGTG GACCTGGCTG
181 CTGAGGAGAC ACCTGCATCT GACTAAGAAG AGCCACGGTG TTAGTTGAAT GGTGTGGAGT
241 AGGGTGGTTC TGTGGGACAG TAGAAAATCG AGAGGCATGT GCCGTTAGT GAAGTATGAG
301 AAGCTACCCC AAACGACAGA GATTGTCACT CAGGCCAATC CGTTTCGAGT TTGATGGGCA
361 GCCGGACAGT GAGACAGACA CACCTACTCA GTTGGAGGAA GGATGAGAAC AATGGCCAGC
421 AGGGATTGAG AGACCCTGAC AGGCGCAAGG CCCTAACACA CACACCTACC ACCTCACTG
481 ACAAAGCTGC CAAAGACCAA AGACTTGTTC TCCATTAGAA ATGACAGCTG GCTTGACCCG
541 ACAGCATAAT AAGCAGAGTG TACTCTGATT GGAGAACTTT AATGTGTTTC ATTCAGTATT
601 ATAAAAGGAC AGTATTACAG ATTTTGTGTT ACACCTGCTG TACATGTGGG GCAGTGTGTC
661 TTAAAGTAGG GTAAAGTACT CTTTAAAAAT GGGTCCTAGA TATTTTTTCC TTAACTCAA
721 GTCTCTTACT GTTTAAATGA TTTTATTTT GTTTAATATG GAGGAAAAAG AAGCGTAAAT
781 GGACAATATA TATTTAGAGA AAGATGGTTA GCTGTCAGAA AAATATGCAA ATCAAATCA
841 CACCAAGACT GCAGCACACC CCTGTCAGAT GGCTGTGATC AAGAAAATAA ATGACAATGA
901 GTGGTGGTGA AGATGTACTA AAGGGAAACA CACACACACA CACACACACA CACACACACA
961 CACACTGGAG CAACCACTGT GGAATCAAGT ATGAATGGTC CTCAAAAACC TGAAGATAGA
1021 GCGGGGCGTG GTGGCATAACA CTTTATTCC CAGCACTGGG GAGGCAGAGG CAGGTGGATC
1081 TCTGAGTTCC AGGCCAGCCT GGTCTATAGC ACAGGTCTA GGACAGCCAG GGCTACACAG
1141 AAAAAACCTG CCTTGATTAA ACCAAACCAA ACCAAACCAA ACCAAACCAA ACCAAACCAA
1201 ACCAAACCAA ACCAAACCA ACCAAACCAA AACACTGAAG ATAGAAGTTC AGTATTCCAT
1261 TCCTAGATAT ATACCCAATG GAGACTAAGT CAGCAAGACA CCTGCACAGC CATGTTCACT
1321 ACTACACTGT TCACCACAGC CAGGCTGTGG AACAGCCTG AGTGTCCATG ATAAATGAAT
1381 GGATAGGTAA CTTTCAAGGT AAATGGACTC TGCTGTGTAC ATGCCTCACA TTCTGTTTAT
1441 TCATTTTCT TTAGAGGTG TCCATTCAGG AGTCACATGG TAGTTCTATT TTCAGTCTTC
1501 TGAAGATACT AACTGGTCC CCACAGTTTA CACTTTTATC AGCAGTGAAT AAGGGTTCCT
1561 CTATCCTTAC CATCATTTGT TGTAAATTTT CTTGATGACC CTCTTTCTGA CAGGGATAGG
1621 ATGTAATATC AGTGTGAGGA AGTACAAGT GTTTTCTAAG TATTTATTGG CCCCTTGCAT
1681 TTCTTCTTTT GAAAAGTGTG GGTTCCTGAC ATCTGCTCAG GTATTCAATG GATGTTGTTT

```

Figure 5 (continued):

```

1741 CTTTGGTGTT TGAGTTCTTA TGAATTCTAG ATGTTAAATC CCTGCCTGTG GTTCTCTCCC
1801 ATTCTGTAGG CTGCCTCCTC ACCCTGGCAA TTGTTGTCTT TGTTTTGCAG AAACCTTTGA
1861 CTTTCATGGAA TCTCATTGTG CAGTTTTCCC TCCTCTGCTA TAGCCTGAGC TAATGCACTG
1921 GTTTTTACAG AGCCCTGGTC TATGCCTTTA TCCTCCTCTG GCAGCTTCGG AGTTTCATTT
1981 CTTACATTTA GATCTTTGAT CCACCTTGAA CAAGTTTGG AGCAGGGTGA GAGATACGAA
2041 TCTAGTTCCA TTCTTCCATA TGTGATCCTA GTTTACATAG CATCGTTGGT TGAAGAGGTT
2101 TTATTTTATT TTTAAATAAT GTGTCATAAA AAACGAGGTG GTTGTAGCAG TGTGGATTG
2161 TTTCTTTGTC CTTTGATCTA CAGGTCTTGT TTTGTGTGAG TCTCATGATG TTTTATTGCT
2221 ATGGCTCTGT CATAAGTCTT GAGGTGAGGT ATTGTGATAT ACCTTCAGTA TTGCTCCCTC
2281 AGACTCAGGT TTGCTTTGGC CAGGAGTCAT CTTACTCAGT GCTCTTAGAG CTCCCCCAGC
2341 ATGTAGCTGC TACTATTCTT AGTTGATAAA TCAGGAAACT GGGGCTCAGA GAGATTAAGT
2401 GTCTTGAATC ACTTCTGGGG AGGTGAAACG TGGAGACACT AAAGTGTGTT TACCCTGTAC
2461 TGCTCCAGTA GCTGTCGGGT GCTGGGCTAC AGCAAAGCAC CTATACTATA TATTACTCAG
2521 GAGGTGGAAA AACTCAGCCT CCCTTGGGGT TCCCAAGCTC CCAGGTGTCC AGTCACTGCT
2581 GGAAACCTCA TGGAGTCTGA AAGGAAGGGT TGAGGGTACA TGGGGCAGCG ATGAGGAGCC
2641 TGGGGCTGGG ATCTCCCAA CACCTGGATA TCCAGATGCC ACTGGGTCAG GGGGAGTTGG
2701 GAACAGAGTT GGGATGTCCA TGGACCTGTG ACAAGGCCAG GGCCAGGGGG AGGATAACTC
2761 TGGCTTTACT AATTTGCGAA AGTCTTAGC TTAGCAGCAG TTGTCTGGGA GCACAGAGGG
2821 GCCTTCTGTA AGAGGCTCAG GCAGTCCGCG TCTGTAGGCG AAGGTCTTCT CCATGTTCCC
2881 CATGGTGGTT CTTGATGAAA GAGACAGTCC TTGGCTCCAA ACTGGTTTAT TGATTGTTCA
2941 TTGTGGAAAA TGGGTGCACA CCACCTTCTC AGGGTGGACC AGAGATCAAA TACCTTTTGC
3001 AGGGAGGAAT ATCTGGGAAG GGACGCTTAC TGGCTAAACC CTCAGGGCCT CTAGATACAT
3061 CATTAGCATG GAGAACTCTG TTCTGGGCTA CATGACCACA GGCCACATT CCACAAGCCA
3121 CATGTGGGAA GTGTGGCACA TGTTCTAGGC CAGGAATCTG GTAGGGAGCG TGGAGCCACC
3181 TACCATCCCA GGTGGGTGCC TGGGTGCCAG GGACCTGAA CCCGCTCAAC CTTACCAAGT
3241 TTCCTGGCAG GGTCCACTGT CCTACACAGA AGCTGGAGGA GGTGTGAGGG TTGTGTCTTT
3301 GTGGAATGTC CCATGCTGCT TGGGGCTCAG TTTCTCCACC TGTACCTCAT TGGTTTGGGT
3361 ATAAAAAGTG GGGATACTTT ATTATTCTCT GACTCGGTCC TGAGGAAAAA GCATCGTGGC
3421 AGTCCAGGAA CCACACCCTG AGGTTCTCTG ACTGAAGGGA CTCCCTAAGT CTCTGGAGTC
3481 TCTCCCTTTC ACAGAGCTGC CAAAGTCTAG GTTCTTTTGA GGATAACAGA GCCATGCTTG
3541 GTAAGCAGAC AACAGCATTT GTTTACTCAA CCTTCTTTTG TCAGCTCCCT CTTCAATAAC
3601 AAGTTGAGAC ACCATGCTGG CTTGAGGAAG ACTTCTAAAG CCAGACAAC GTGCAAGGAA
3661 GAAGAAGAAG GGGCAAGTGG AGTTAGCCTG GATGTAGCCC TCAAAGTCTC CAGAGACCAG
3721 CCATGAAGGC TCAAGTGGAG GGCAAGACCT GCAGCAGCCA AGCATCTGGC AGGAGAGGAT
3781 CCTGGGAACC CCTCTACCAT GACACACATT CTTCTGTCAG GTCACACTTA ATAGGCCATT
3841 TCTTATTTGG ATCTATCATG GTGTTCTGTG CGAGATTAAT GAGGTGTTAT GCTGCGAACA
3901 GAAAGTTATA TAAAAACAAG TCCCCCCCCC TTGTCACTGC TGCTAAGAA GTAGCAGAAA
3961 TTGTCTCAAG TGTCTCTCTA ATCAGAAACA ATAAAGGTCT CCTTGGATT CAGCCCTCCA
4021 GTTTCCTCCT TCCTTGCTGA GCCTTGGACA CCCATACAAA CCTCCTGGAT GCTACAGCTC
4081 TGGGCAGAGA CTCCAAGGTG GGGAGAGACT GATGGTACAA AAGCAAAATA CTTGTTTGGG
4141 GGTACACCCA CTCCTCTGCC TGTGTGGTTC CTGCAGTCAG TCCTGCAGAC AGGCCCTCAG
4201 TGGGTCTTCC ATGGGCAACA CGCAGAGGGA GGCAATGGAT GGGAAATACCC ACACCCTGGT
4261 TAGTTTACCC CGGCCATGCT CTCTGCTCTT CATCCCTCCT CTGCCCTCTG CCACGGCTTT
4321 CTCTGCAGGA ATCATATCTT CATATTGGCC CACAGGTGTT CTCCTCACC TAGCTATGAT
4381 GTTTACTTTA GAGTGACCTT AGCAGGGCTG GTGGGAATGA GTTCTAGAAG GCTCACGGAG
4441 ATGCTAGGGA AGAAACGTCT TCTAACTACT GAGGTACTA AGTTCCTGGT GGTGTCTCT
4501 GCCTTTCCCT TGTTAAAGTC ACCTTGAAGT TAGTGACAG GAAATCAGAG CCCAGTCACA
4561 GAGTAAATAT GGTCTCTGAAG ATTTCTTTTG AGTGCCAGA ATCCATGACA TTTCAAGAGC
4621 CCTCTTTGTA CCTTAAGTCA TTTGGGGTTG TATCTTCTGC TTGATGTATG TGTGTGTGTT
4681 TATCAAAGAG TGAGATGGTT ACATAAGAGG TGCTCTAAAG GACAGAGAGG ATTTGCAATT
4741 GTGGCATGTG ACATCCTCAG GCCTTGCTCT GGTGCCAGGA GGAAGTATG CAGAAAAGAG
4801 TAAGAGGTCA TTTCTGGAG GCTGTCACTA TAGAGGAGAT CTTACAGTGC ATTCCCTCCT
4861 CCAGGCCCTG CCTGAGGATA GACATGTGCT GACTGCAACT GAAACAGAGG CTTGGGATGG
4921 AGAGTTAGGT TCACAGAAGG GAGGGTGGGA GATGGATGCT TGCTGGGTTT TGGGTCTCAT
4981 CACCAGCTCC TGACCACCCG GTCAGCCCAT GTGCTTATTC CATAGCTTTC TTTTGCTATG
5041 TTTACTCAGT GTGGTGTGTT TTGGGACCCA GCAGAAAGCCA GTCCCAGGCT GACAGCTGTG
5101 GATACACAGG GCAGCATGAG GGTCTCTCAG CTGAAGCAGT CAGGCTGGCA GAGAGAGAA
5161 ACCAGCACAC ATTCTTCAA CCAACTATGT CTTGAAAAAC AAACATATTA TATCACATAT
5221 ATTGCATTTA TGAGACAGCT AAAATGTACT CGGGTAGCAT GACTCCAGT GGGGATATCT
5281 GCAAGTGCCA TGAGTGGCAG AGGGACAGCC AATGTGAGGC AAGAAGGAAT TCTGGCTCAA
5341 CACAGCTTAG CTCCCTGGTG TTGGTTCAAA CTTTGAGAGT TTGACCACAA GCACCTTATT
5401 TTTGACATAT TTAAACAGAG CACAACTTTG GGAAAAAGTT TTCTTATGAA AATTATCACA
5461 ATAAAGCTTA AGGCATGACT ACATTAATAAT GCCTTTGCAA AGTATATGTG CCCTCTTCCA
5521 CAAGAATGGT TCTATTGACT GAGAAATAAT GTTCAGGATA AAGATCCAGG AAGAAAAGAT
5581 CAGGSATAAG TAAAAATACTA AACTCTTTTG CAAAGTACAT AGACCCTCTT TCATAACAAT

```

Figure 5 (continued):

```

5641 GGGTTCTATT GACTGACAAG CACTGCTCAG GAGTTGGGAA AGAGTCTAGC ATAAGCACGA
5701 TAGCCTGGAG ACTCTAGTGA GGTCTAGTCT TACAGACAGC AAAAATCACC AGGTTACAAA
5761 CTACATTCAT TTCCAGTTTT CTGATCAGGC ACAGGTATGA ATCCCTTCTG TTGAAGAGAA
5821 AAGTCCATGT GTTTAAAATA TCTGGTTTCT CCAGTGCTAT TAGCGAGAAG ACTTGAGCCC
5881 TATACAACCT CCACCTGGAG TGACATCCTG TCTTCATGGT ATATTACATA CCTAGACACG
5941 CTCATCTCAC AGACTTAGGA CTTTGTCTTC TGATCTCCAT TTCTGATCCC ACTTCCACCT
6001 TTGCCTTGAT AGTGTCATTT TCTTCACTGC CTTGGTGACA ACCATGTTAT CCTCTGTGTA
6061 TTTGAGTGTT ACCATTTTCA GATTTTACCT GTATGCAAGA TCACACAGTC TTTGTCTTTC
6121 TGTCTGGATG CATGCTAATC TCTACACAAC AACCCTTCCC CGTCACTCAG ATCTTCTCTC
6181 ATTAACACAT ACATGGTGCT GAAGAGGCTA GGGAGCTTCC CTTCAGTGGG GAGCTAGCTG
6241 GCTATTGGGC CTTTTTGACT GTCCAGGAAG GCCCCCAATT GCTGAGACAA GAACCTTAGAT
6301 TCTTCATTAT TGACTCTAAC TCATGTATCA AGCAGAAGCT AATGAATAGT TATCAACAGG
6361 ATCAGAGGTT CCAGTGTAAG ACACCTTGAC ATGAAAGAAC GGAGGAAGGA CAGATGGATG
6421 CATAAAAGCA GGACCACTGC CCCAGGAAGG TCCTGGAAAC TGATGCAGGG CAAAGGACAG
6481 GTTATAAACC AAATCTTAGG GAGTCAGGAA GAGCACAGAG GAGCTCAACC AACTGACCAC
6541 TGCTTAGGGG CTACCAACCC AATCCTCCCT GTGGGAACAG CTAAGCTATC AGCCAAGGGT
6601 AATAAACAGG CAGGACCTGT GGATGACATG GAGAGCATAG GGACCCTGGG TCCAGCCTTT
6661 AGCACCTGCA CTCTCAGGAT ACTCCACCAT TGTGTCTTAG AGAGCCTAGG GATACTGGGT
6721 CCAGCCTTTG GTACCTTCAC TCTCAGGGTA CCCCATCACT GTGTCTTGA GAGCCTAGGC
6781 ACCCTGGGTC CAGCCTTCAG TACCTGCGCT CTCAGGACAC CCCACCATTG TCTCTTGCCC
6841 CGTCTCTTCT TCCTCTTCTT CCCTTTCATT GTCTCTTCTC TGTTTCTTTC TTGACTCTCC
6901 TTTCCCTCA CACCCTCACT CTAGTTCTCC CCTTCCCTCT CTGCATCACC CTATTCTCTC
6961 TGTGTTCCCT CCACCTTCTT TTATCTCTCA TGCTTCTCTC CTCCCTCAA TACTTGTAC
7021 CCACATACT TCAGGGGCCA GCTCTAGTGA CAAAGCTGTT AATAGCAAGA CTCTCAGATC
7081 TCCAACGGCT CAGAGGAGCC AGACCCACCA AGAATCTCT CCAGGTCCAA TTTGAGTTT
7141 CTTGAAAGC TTTGAGCAA TGCTCAGGGA ACATGCCACT AACAAGAAGA TGCAAATTCC
7201 AGTTGAGAGT GGGAAAGGCC CTTGCGTAGG TCCCATCTTC CAGGCCAAGG TCAGAGGGGC
7261 TCTGTGTAAT CCGGATTGAC AGGGCTCAGA ACAATGTTTT GTTTTTAAGG TTTATTATT
7321 TTAGGTGTTA GTGTCTTTCG TTGCATGACC TTATGTGCAT CATGTGTGTG CAGGTTCTTG
7381 ATGACAGTAG AGGAGGGCTT TGAATCCCTG GGGATAGGAA GTTACAGGAA ATTATAAGCT
7441 GCTTTGTGGG TCTTCTAGCT TTCCCAACAG AAGTGAATGC TCTTACCAC TGAGCCATCT
7501 CTCTAGGCCC AAGAGACATT GCTTTATGGA TATAATTGTG TGTGTGTGTC AACATTGAGG
7561 AAAGGGAAAT AAAAAAAAAA CTTCAGCCCG TAAGGTTGTA CAGTTTCACT AATTGCTTACT
7621 TTTAGTTGTG ATAAAAATGGC AGGTGCTTCA ACATTTATAT ATACAAAAAC TTCCCTGCTG
7681 GTGGTTCAAC TGTGAGAACT GGGGTAAGTG GGTGAGTTCT CTTTTTCTGT CTCTGTCTCT
7741 GTCTCTCTCC TTCCATTCTT TCTTAAAGGA AATAAACATT GCAGCTGGGT TATAGCTCAT
7801 CAATATGGAA GTTACAGAAG TGAAAAAAGG CATTGCCTTG GTGGGTGGTG TTACCAGCTG
7861 ATTTTGTGGT GTCCTGCAAG GAGGTCTGGG GACTGGCTGC TCTGTCTCTG TCTGTATGAG
7921 TGAGGGAAGT CTGGGGAGCA GATTCCCTAA CCTTCAGCCT CCGCTGGTTC CTGAGTGAAC
7981 CCAGCCTCTC TGGTCCTAGT AGCTTTTTTC AAACAGGAAT CTGAGTGGTG ACAGGGAACA
8041 AGTACCAGCC CATTGCTTAA GTGCCAGGGT TAGTGAGGGC AGGAAGCTGC CATAGCTGGG
8101 ATTAGTAGTT GTATTGGATG TAGGAAGTCC TATCTGGGA CAGCTAATCC TTAATGCTTC
8161 ACTGGAGATT TTCAATGAGA AATTTATCCC ACGGCCATA TGGCCCCATC CTTTTGTCTC
8221 CAACAGCCAA GTATTTTCCA TTAGAGGAGA CTTCTGTGAC ACTTGATGGA TGCTATTCC
8281 AAGGTGACTT GGGGCACTCA GTACAGACTT GGGATGACCT CTGACAGCCT AACCTCTCCC
8341 CAACAAGGGC CCTCTATGTT TGCTATGTAA TGTAATGTCA GACATTGTCA GGAGTGTCCG
8401 CAGCACAGCC TGCCAGTGT GAGGGCTCTC ATAGGTTTCC CACTGTCTTA TCTACACAGG
8461 GATAACGAGG AGGTAAGCTG CAGTTCCTAG TCTCACTTCA CAGAGGAAGA GATAACCCCA
8521 TCCCAGGTCA TGTAGCCAGC AGTGGAAGA ATGAGGATTT GAAGTCAAGT CTTCCAAGTC
8581 CCATTGATAG CATCTCCTCA CAAGTCCCTT GCCACCCTCA CGATGCCTTA GACACTTGCC
8641 TGCCCTTTAT ACTAAGGAGA TGCAGGTACA AGGGGTTTAC CCATGTAGCA GCTGAGGCAG
8701 CTGGGGATAG ATACCAGCAG CAGGCTGAT GTCACCACTC TAACTCCAGC ATCCCCAGTC
8761 TGTGTTCCCTG GAGTGTGAAA ATCCCTACTT AACAAGATTG TGCAACAGTC CTTGGCTCTG
8821 TGACCCATAG CTGGAACAG GATTCTCATT GATTGTGGA ACATGGTGGC AGCCAGCCAA
8881 AAAGAGGGTC TGCATACAGA AGACAGTGT GGCAAGGCCA CAGCAGACTC TGACTACCTT
8941 AGCTTACAGA ATTACAAGGT CATAATGTCC TCTGCTTTGG TCACCTCATG TTAAGGACAG
9001 GCCCTAATGA AGATGGGGCA GAAGACTGAA GGAATGGCCA ACCAATAACT GGCCCAACTT
9061 GAGACCCATC CTACAGGCAA GCATCAATTC CTGACACTAC TAATGATACT CTGTTATGCT
9121 TGCAGACAGA AGCCTAGCAT AACTATCCTC CGAGAGGTCC ACCCAGCAAC TGACTGAAAC
9181 AGAAAAAGAT ATCCACAGGC AAACAGTGGA TGGAGTTCAG GGACTATTAT GGGAGAGCTG
9241 TGGGAAGGAT TAAAAACCTT GAAGGGGATA GGAACCCAC AGGAAGACCA ACAGAGTCAA
9301 CTAAGAGACC TGTGGGAGCT CTCAGAGACT GAGCCACCAA CCAAAGAGCA TACACAGGCC
9361 GGTCCGAGGC ACCTGGCAGG TGTGAAGCAG ACATGCAGCT CAGTCTCCAT GTAGGTCCCTC
9421 CAATAAGCGG TAGCCTGACT GCAGTATCCA ATCCCAACA GGGCTGCATA GTCTGGCCTC
9481 AGTGGGGGAG GATGCCCTTA ATCCTGCAGA GACTTGATGA GTGGAGAGCT ATCCAGGGGG

```

Figure 5 (continued):

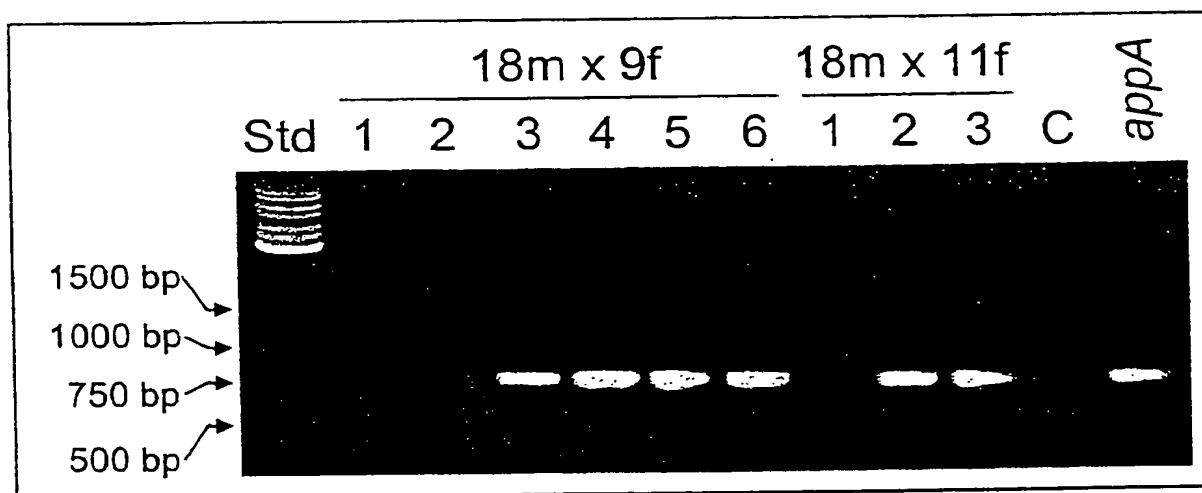
9541	AACCCACCCT	CTCTGAGAAG	GGAATGGGGA	TGGGGGAGGG	ACTCTGTGAA	GAGGGGACAA
9601	GGACAAACAA	GAACCTCAAA	TAGGTCAGGC	CCTAAAGGCT	TGCTAAGTAG	CAGTGGCCCA
9661	GCTCTGTCCT	GTTCCCTCAGC	CCAAGGCTCA	GCTCCACCT	GTTTCTGTGT	TTTTCTGGCT
9721	TTTCATGGGC	CTAGGACTTG	GTGACCAGTT	CAAACAATGG	GGCCTGTGGA	AGACACAATA
9781	TACAAGACTA	GGGACATTCC	TGTTCTGCTG	ACTATCCATA	GCCTGATGTA	GGTGGGAAGGA
9841	CCCAATCACT	GGATTTCTAC	CCTTGACAGC	CCTTGACAGC	TGAGGGCCCTC	TCAGAAACCT
9901	ATTTCTTCCA	CTGAAAAATG	AGACTCTCAA	ATGAACGTCG	TGACAATCAT	CAGGCTTATT
9961	AAAGAGGTGT	ATCTAACCTG	AATGGCAAGC	AGACAGCAGG	CAAATGTCTG	TATCAACCTC
10021	TAGGAAGGAC	AAGAACTGCT	CACTGCTGCC	CCCCAGGAGG	CCATTTGCTG	AAACAGCTGC
10081	TCTCCTGCTG	GTGCACAGGC	CCTGCCTTCT	CATTGCAGCC	ACAGCCCCTT	CCTGTCTGAA
10141	CCTCCTGTCA	GGTCACTGGG	AAACAGATCA	AGATGGAACA	GGACAGCTCC	TGATGGTAAA
10201	TAAAAAACAG	TGGTCATGGC	TATTCATAGG	GGTTTATGCT	TCTTCAGTCC	ACACTGTGAA
10261	GAGCTGTGGG	CATGAACCAC	AGTGTTTCGAG	GTAGAGTTGG	GGTTCTGAAA	TTACAGTGGG
10321	GGTGAGCTCA	GTAAATGTGA	GCTGGAGGTC	ACTCGTGAGA	CACACAGTCC	TGCTGCTTCT
10381	GTTCCCAATA	TCCTGAGGAG	ACGACACATC	TACTTTGTTC	AGAGGCCACA	GTCTAGTTGA
10441	CCTGAGAGTT	ACCAGTTTCT	TATTTGTGTG	TGTGTGTGTG	TGTGTGTGTG	TGTGTGTGTG
10501	TGTGTTCGT	GTGTGAGTGC	AGGTGCACAT	ATGATAGCGT	ACACGTTGAG	GTGAGGAT
10561	AACATACAGG	CGTTGTCCCC	TCCTACTTTT	CCTCGGACTC	TGGAGAACAA	ACATGGGTCC
10621	TTATTCCAGG	GGAGCAAGTC	GCTGTGGCT	GACACATCTT	GCTCACATAC	ATTTTACCTA
10681	GACAATGGAG	CCTCCATCAG	AGTATTACTT	TAGCTCCTCA	CCGATGGCAA	TGACCCACCT
10741	CTCTACCCAC	ATAGGAGTTG	GGTCTCCACA	CACCCCCACA	CCCCCTTCAC	CAAAACGTTT
10801	TCAGTTACTT	TATCTGGTAA	AGTTCATCAG	AGAATGAAGC	CAGTATTAAG	AACATGGAAT
10861	CATTTGGGAA	CCTGGATCTA	GCAATACCCC	ACCCTAGATG	GAGTTGCTGA	GTTTTACCCT
10921	CAGATTATAA	TTCCCCCCTA	GCTTCTATGG	TTTATTCTGA	AACCAGGGGA	ACTCCGATTCC
10981	TCCCTTTGGA	CCACAGACAT	CCTGGCTTGT	GAATTCACAT	GTCATCTACT	GCTAATCCAT
11041	TGGTAGTATG	TGGCTCACAG	AGACACACTA	CAGTCATGGC	CAATGTCAAG	GTAGGACAGA
11101	TGTGAATCAT	TCCCCCAGTC	CTGCTGTTTT	CATGACTAAC	CCTCCTCAGC	ACAGTGACCA
11161	TGAACCTACT	TTTCCCCCTC	TTTTATTTTT	AGAATTGCTG	GAATTTTCTA	TTTTGAGAAA
11221	TAATAGCCTT	GGGCAGCATT	AAACAAAATC	ATCTAGAAAG	CTGGTTTAAA	ATACAGATGG
11281	TTGAGTCAGT	GAAAGAGTGA	GGAATGTCAT	TATTGGCCCC	TCACAGAGGC	TGGCTCACTC
11341	CAGCAGAGGT	GGTTGAAGCT	CTTGGACACG	GGTCAGGTGC	ATAGGAAAGG	TNGTCTGGGA
11401	CACTGAGAAC	CACAATTGAA	CAAACAGAAC	TGTTGGCTTT	TTTTTTTTTA	AATGAGTTCT
11461	CAAAAAATGA	CTGGCTAGCT	TAGGCAAATA	CTTCGAGCCA	ACCCAACAGA	ACATTCTTCC
11521	ATTGATTCAT	TCTGGATCTT	CTTCTAGAC	AATACTGAAC	TGACCCCTTG	TTGGCAGTCT
11581	CAAGTTTGAC	AACATAGGGC	TTTGAACCTG	GCACAAGGTC	CATCACTGTC	ACCCAAGCAT
11641	CCTGGGTGAC	CTTTGGGTG	GAATATCTTG	GCTAACCTTA	GATATTTTCT	TTGGAGTATC
11701	TTTAGAACAT	CCAGGAAATA	GGGCTTGATT	CTCATCCTGG	GACCACAATA	TAAGTACCCC
11761	TAGAATCCCA	GGAGATCGTG	CAGAGAAACA	AGGATCTCTC	TCGTGTGCAT	CCTTCTCAA
11821	AGCAGTGAGT	AGTGACTCCA	CTAAACTGAG	TTCCCATCTG	AGAGTCCACA	GGAGGCTTGT
11881	GGGCAAGAAG	CAGAGGGAAG	GCACCTGTTT	TGTTGGTAAA	GTTTTGACTC	TAACAAATTT
11941	GAAGACATAG	ATGACATTGT	GTCAGACTAA	CAACAACCTA	GACTCATGTG	GGTTCTGTGT
12001	AGGGATCAGA	TTTTATTTCAT	CAATGACTTG	TCTTAGTGTA	TAGAGAAAGG	CTTCCTACTG
12061	GAGTGTAGGC	TCAATAATGA	CAGAAGAGAT	AGCTATTTCC	CCTAGGGACT	GTGCTGCTCC
12121	AAGTTTGGTG	GAGAAAGGCA	TGTGGGAACC	TAGATGTGCT	CTCTGGGGAG	GGGGTCTGAA
12181	GCTGGCTTCA	TAGAAGGTGT	GAAGTTTTCG	TGAAACATCT	AAACAGAATT	ATAGCTTAGG
12241	AAAGTGAGCA	GGCAAGGCAG	GGAATGTGTT	GCATATGTAT	ATGTACATGA	ATATATTATG
12301	TTATAGATAC	ACACACATTT	GAACCTCATT	TGCAGATGAC	AGAAAAATAGG	TTATTTTGCC
12361	TCTCTTAAC	GCTAAGCACA	ATGACTTCCA	GTTCCATCCA	TTTCCTGAAA	TGCCACAATT
12421	TCATTTTTC	TTGTGGCTGA	ATAAAATTC	ATTGCAGACT	GGGCCCTACT	TCATCCACTC
12481	CTGAGGGCAG	GCATATCCCC	TGGCTCCATT	TCTTACCTAT	TGTGAAGAGA	AGTGCAACTG
12541	TCTTGTGAA	AGGCAAGCGT	GAGAGAGGCA	GGCACTAATT	GTGGGTTTT	GTTTCTTCTT
12601	CCTGCTATGA	CTCTCCATTT	GTCAGAACCA	AAGATCGATA	AAAGCCGCCA	CCATGAAAGC
12661	CATCTTAATC	CCATTTTAT	CTCTTCTGAT	TCCGTAAACC	CCGCAATCTG	CATTTCGCTCA
12721	GAGTGAGCCG	GAGCTGAAGC	TGGAAAGTGT	GGTGATTGTC	AGTCGTCAATG	TGTTGCGTGC
12781	TCCAACCAAG	GCCACGCAAC	TGATCGAGGA	TGTCACCCCA	GACGCATGGC	CAACCTGGCC
12841	GGTAAACTG	GGTTGGCTGA	CACCGCGCGG	TGGTGAGCTA	ATCGCCTATC	TGGGACATTA
12901	CCAACGCCAG	CGTCTGGTAG	CCGACGGATT	GCTGGCGAAA	AAGGGCTGCC	CGCAGTCTGG
12961	TCAGGTCGCG	ATTATTGCTG	ATGTCGACGA	GCGTACCCGT	AAAACAGGCG	AAGCCTTCGC
13021	GCCTGGGCTG	GCACCTGACT	GTGCAATAAC	CGTACATACC	CAGGCAGATA	CGTCCAGTCC
13081	CGATCCGTTA	TTAATCCTC	TAAAACTGG	CGTTTGCCAA	CTGGATAACG	CGAACGTGAC
13141	TGACGCGATC	CTCAGCAGGG	CAGGAGGGTC	AATTGCTGAC	TTTACCGGGC	ATCGGCAAAC
13201	GGCGTTTCGC	GAACTGGAAC	GGGTGCTTAA	TTTCCGCAA	TCAAACCTGT	GCCTTAAACG
13261	TGAGAAACAG	GACGAAAGCT	GTTTCATTAAC	GCAGGCATTA	CCATCGGAAC	TCAAGGTGAG
13321	CGCCGACAAT	GTCTCATTA	CCGGTGCGST	AAGCCTCGCA	TCAATGCTGA	CGGAGATATT
13381	TCTCCTGCAA	CAAGCACAGG	GAATGCCCGA	GCCGGGTG	GGAAGSATCA	CCGATTACCA

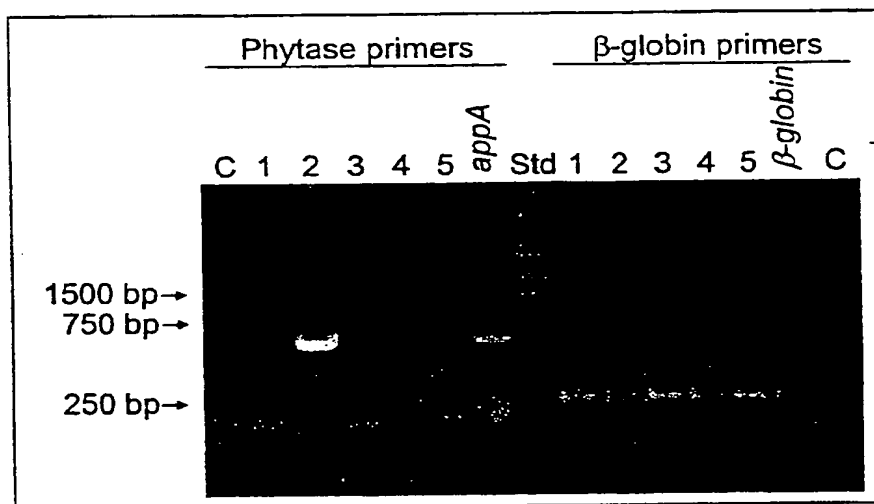
Figure 5 (continued):

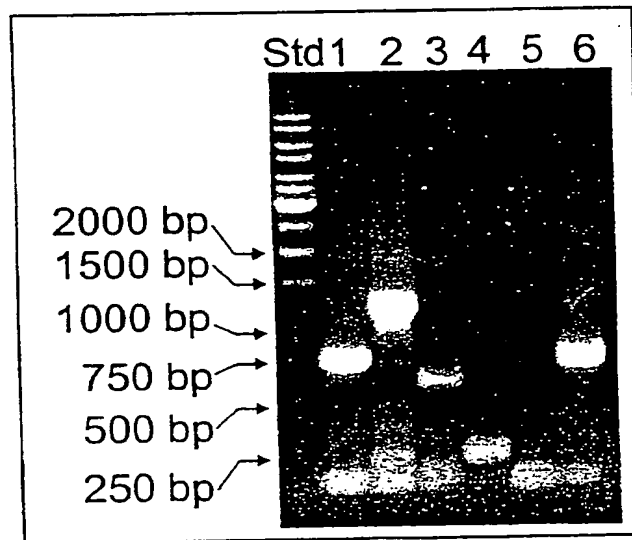
13441	CCAGTGGAA	ACCTTGCTAA	GTTTGCATAA	CGCGCAATTT	TATTTGCTAC	AACGCACGCC
13501	AGAGGTTGCC	CGCAGCCGCG	CCACCCCGTT	ATTAGATTTG	ATCAAGACAG	CGTTGACGCC
13561	CCATCCACCG	CAAAAACAGG	CGTATGGTGT	GACATTACCC	ACTTCAGTGC	TGTTTATCGC
13621	CGGACACGAT	ACTAATCTGG	CAAATCTCGG	CGGCGCACTG	GAGCTCAACT	GGACGCTTCC
13681	CGGTGAGCCG	GATAACACGC	CGCCAGGTGG	TGAACTGGTG	TTTGAACGCT	GGCGTCGGCT
13741	AAGCGATAAC	AGCCAGTGGA	TTCAGGTTTC	GCTGGTCTTC	CAGACTTTAC	AGCAGTGGC
13801	TGATAAAACG	CCGCTGTCAT	TAAATACGCC	GCCCAGAGAG	GTGAAACTGA	CCCTGGCAGG
13861	ATGTGAAGAG	CGAAATGCGC	AGGGCATGTG	TTCGTTGGCA	GGTTTTACGC	AAATCGTGAA
13921	TGAAGCACGC	ATACCCGCTT	GCAGTTTGTA	AGGTACCCGG	GGATCACAAC	TTGCCCTCTG
13981	AAGAGGAAGA	ACAGAAGGAT	GCCACAATCT	TCCTGCTGGC	TACTCTCCAG	TGGTTTCATC
14041	TTACTTCTGA	TGGCATTTC	CTCTAGAAAG	TGCTACTATC	ATCCACACAT	TTCTACCTGA
14101	GACCACCCAA	AGGACCCTCC	CAAATTCTCT	TCCTCTCTGA	GTAGTCTCCA	CACCTGTTAC
14161	CACCATCCCA	GAATTAAAT	CCTAAGTCCA	CTCTGGCGTG	TGACTTGCCT	CAGTCTTTCG
14221	AATAAGAGTT	GTTGGCAGTG	CCAGGCGTGG	TGGCGCACGC	CTTTAATTCC	AGCACTTGGG
14281	AGGCAGAGGC	AGGCGGATT	CTGAGTTCGA	GGCCAGCCTG	GTCTACAGAG	TGAGTTCACG
14341	GACAGCCAGG	GCTATACAGA	GAAACCCTGT	GTCGAAAAAC	CAAAAAAAGT	AAAAAAGTT
14401	GTTGGCAGAG	TGTGGGTTAT	ATACCAAGTG	GAGATTTCAC	ATGAGTGGCT	GAAGCTGTAG
14461	CCAGAAGGAA	CTTAGAGGAT	AGCTCATAAC	TTAAAAAGAA	ATGTAGAGAG	TAGCAGAAAC
14521	ATTGAGAGAG	TGGGCACACA	GCCACTGTGT	GAATGTGGCA	GAACACAATC	CAGCCAGCTA
14581	TACATGCATA	AGTGTATATT	GGCGCCATCC	TGACTGATGA	GACACAGGAA	AACAGATAGA
14641	CGGGGTTAGG	TGGCCATGGC	CTTTCCTGCC	TGCCTCTTCC	TAAGGGTCAT	CTCAAGACCT
14701	TATGCTCTCT	TAACCTCTCC	ATTGCTACTT	AGCTTCTAGA	TATCACCTCC	AGATTAGTCT
14761	CCTTGGGTAC	ATCAGTGATC	CTGGTGATAT	CCAGGGCTTC	CTGATTCCAT	CTTTGTCTATA
14821	GAGGCTGCAA	CTAAAGAGGT	CTTCTTAATA	CTTCACACCC	TGATGCCAAA	AGGAAGACAC
14881	AGAAGTTCAC	AGAGGTGAAG	TGATTTCATGT	AGGACATACA	GTGAGCAAGC	ATCAGGGTCC
14941	GGATTATCTG	ACTCTACTCT	AACCTTTATG	TAAATGTGCT	TTATGCCATT	AACACTGTCA
15001	TTCTGTGCT	TCAGCTCTGG	GAGACTCCCA	AGCACTCTTA	GGCACAAGCC	ACAATTAAGG
15061	GACTCTGACA	CTCTGCATTG	ATTAATTAGC	ATGGTGGTCT	CTATGTTTCC	GAGCTCATGA
15121	TTGTTTCACT	TTCCATATAG	GCTATGAAGG	GTGTGAGGAA	ATTTTTTGGG	GACAGAATTG
15181	GAGGCAATCC	ACCTCTCTCA	GGAAGCCTCT	ATCTGGAAAA	GCTTACAACCT	CAGGGACAGT
15241	AACTGTAGGC	CCAGTCCTTG	GTGTCCAAAA	TGGGTTTTAT	GGTTTGAATC	TGCAAAGCCT
15301	TCCATGTGCT	CAAAGGTTTG	AACATGGAGC	CTCCTCCTGG	TAACACTGTA	TGGAGGCTT
15361	TTGAGACTGG	ATGCTCTTTG	TGCCCATGTT	TTGCTACATC	ATCTGTCAAG	ATATGACCCA
15421	GGCATGCTAC	CAGCTACCAC	AGACTATGCC	TCTCCAGCTT	TCATGTTCTC	CCCACCATGA
15481	TAGACTTGTA	TCTCCTAAAA	ATGGAATCAA	AGCAAACCTT	TCCTGCATTA	AGTTTTTTTT
15541	TTTCTGTTAA	GTGTTTGGTC	ACAGGGACAA	GAAAAACATC	AATACAGATA	ATTAGTACCA
15601	GAGTTGAGGT	TCATTGCTCT	AGCAAGTTGG	ATCAAATTTT	TAGGGCTTTG	GAAGTATTTT
15661	ATAAGAGACA	TGTAGAAGAG	TCTGAAGTCT	TGGGCTACAG	AAGTGTCAAC	AGTTTTTAAG
15721	AATAGTTTTAA	TACACCATGG	GAATTGTGAA	AATCAGAATG	CTCACACAAA	GGCAGACAGG
15781	AAAACGTGAG	CATGTGGCGT	GTGAGAGGGC	ATAAGAAGGA	ACCTAGGGGG	AAATGAGCTA
15841	GAAGCCATTG	GGCTACGTTA	GGGAACGTGT	GTGGCTGTGC	TTGGCCCATG	CCCTGGCAAT
15901	CTGAATGAGG	CCAAATTTTA	AAGGAGTGGA	CTAACTCGAT	TGTCAGAGAA	AATATCAAGA
15961	CAGACCACCA	CTCAGSCTAT	GCCGTGTTTG	TGACCGACCA	GCTACTCTTA	GCCAGCTCTA
16021	TTGTGAAATT	CCAGAGCAAT	TATCAGAGCA	TGAAGATACA	TACAGTTTAG	TGAAGTAAGG
16081	GGTGTGGGTC	CCTAAGTGGA	TGGTGCATAA	ATCTATGTAG	GTGATGCCTA	AGTGACACTT
16141	GATAATCCAA	AATATCAGCA	ATGTGGAATG	TCTTCCAAGG	AGACCTGTAG	ACACACATTT
16201	TAGAACTTTG	CTCATGGCTG	TAATAAATAG	CTAGCTAGAA	ATCATTTTCT	GAAGAGGTTA
16261	GTCTGAGTTA	CGGTTCCAGG	GCAAAACATT	AGTGATGGCA	AGGAAGGCAT	TGCAGTCAGG
16321	AGCCAAAGGT	CAGCTGGTCA	CATTGCATCA	AGAGTAGAGA	GTCAGAGTGT	GAGTAGAAAG
16381	AGGATACAGG	TTATAAAACC	TCACTGTCCA	CTCTCAGCAA	TCCATTTTCT	CCTAAAAGGC
16441	TTTACCTTCT	AAAGATTTTA	GTCTTCAAAA	CCAGTACCAG	TAGCCTGGGA	ACAAAAGTTG
16501	AAACAAATGA	GCCTTTGTGG	GGCATTTTCA	ACTTAAAAACA	GGGCATCACC	TAGGAGGAGC
16561	CCTGTGTGCA	GTAGGAAGTG	TGGCCTCTGT	GTCAGGAATG	CTCAGGCTAA	TAAGGGGTCC
16621	TCTATCTGAG	GGACCCATG	AAGATTCAAC	AAGTAGTTGT	GAGAAATCCC	TGTAAATGGA
16681	TGCTACCAAT	TTGACATTTG	TAGACCTGCT	ATTGTGTGCT	TCTTTATTGG	GCTCTCCCAT
16741	CTCCCAACTT	TCCAACCCAT	ATTCCACATT	AATCCCTTCC	ACCACCATGC	AACACTAGGT
16801	AGGAGAGAAG	GAAGGTTAGA	AGAGAAAGTG	GGTATAGATC	TATTTAGACT	ACTTCTGTCT
16861	GATTAGGGGC	AAGTCCAATC	GTCATTGTCA	GGATACCTCC	AACCAGCAAC	CAGCAAAACC
16921	GCAAAATCAG	AACAGCAAAA	GCAGCCAAAC	AGGCAGCACT	AACCAGCAGG	ATTGGGGTCC
16981	GTAGCGTGGG	AGCAGTCACT	ACTGGTCTTC	TCATGGCTTT	GGCATTAAATA	CTCTCTCAAG
17041	AAATTCCGTA	ATTTTTTTCC	CACCACCTGA	AATTCCGTAA	TTTTAAATGC	AAACTATCTA
17101	CAGCTGGCAA	AAATCACATC	TCTCCTAGAG	CACAAGACAA	ATCATAGTTA	CTGGCTATTT
17161	GCAATCTGAA	GCATCTCAAT	ATCCACACAC	TGGGATTAAA	ACAAAACATC	ATTCATCTTA
17221	CATAACTGTT	TTTTTTTTCC	AATTTTTTAT	TAGGTATTTT	CTTTATTTAC	ATTTCAAATG
17281	CTATCCCGAA	AGTCCCCTAT	ACCTCCCAC	CTCCCTGCTC	CCCTACACAC	CCACTCCCAC

Figure 5 (continued):

17341	TTTTTGACCC	TGGAGTTCCC	CGGTACTGGG	GCATATAAAG	TTTGCAAGAC	CAAGGGGCCCT
17401	CTCTTCCAG	TGATGGCCGA	CTAAGCCATC	TTCTGCTACA	TATGCAGATA	GAGACACGAG
17461	CTCTGGGGGT	ACTAGTTAGT	TCATATTGTT	GTTCCACCTA	TAGGGTCGCA	GACCCCTTCA
17521	GCTCCTTGGG	TACTTTGTCT	AGCTCCTCCA	CTGGGGGCTC	TGTGTTTTAT	CTAATAGATG
17581	ACTGTGAGCA	TCCACTTCTG	TATTTGACAG	GCACTGGCCT	AGCGTCACAT	GAGCCAGCTA
17641	TATCAGGGTC	CTTTCAGCAA	AACCTTGCTG	GCATGTGCAA	TAGTGTCTGC	GTTTGGTGGT
17701	TGATTATGGG	ATGGATCCAC	TAGTTCTAGA	GCGGCCGCCA	CCGCGGTGGA	GCTCCAGCTT
17761	TTGTTCCCTT	TAGTGAGGGT	TAATTGCGCG	CTTGGCGTAA	TCATGGTCAT	AGCTGTTTCC
17821	TGTGTGAAAT	TGTTATCCGC	TCACAATTCC	ACACAACATA	CGAGCCGGAA	GCATAAAGTG
17881	TAAAGCCTGG	GGTGCCTAAT	GAGTGAGCTA	ACTCACATTA	ATTGCGTTGC	GCTCACTGCC
17941	CGCTTTCAG	TCGGGAAACC	TGTCGTGCCA	GCTGCATTAA	TGAATCGGCC	AACGCGCGGG
18001	GAGAGGCGGT	TTGCGTATTG	GGCGCTCTTC	CGCTTCCTCG	CTCACTGACT	CGCTGCGCTC
18061	GGTCGTTCGG	CTGCGGCGAG	CGGTATCAGC	TCACTCAAAG	GCGGTAATAC	GGTTATCCAC
18121	AGAATCAGGG	GATAACGCAG	GAAAGAACAT	GTGAGCAAAA	GGCCAGCAAA	AGGCCAGGAA
18181	CCGTAAAAAG	GCCGCGTTGC	TGGCGTTTTT	CCATAGGCTC	CGCCCCCTG	ACGAGCATCA
18241	CAAAAATCGA	CGCTCAAGTC	AGAGGTGGCG	AAACCCGACA	GGACTATAAA	GATACCAGGC
18301	GTTTCCCCCT	GGAAAGCTCCC	TCGTGCGCTC	TCCTGTTCCG	ACCCTGCCGC	TTACCCGATA
18361	CCTGTCCGCC	TTTCTCCCTT	CGGGAAGCGT	GGCGCTTTCT	CATAGCTCAC	GCTGTAGGTA
18421	TCTCAGTTCG	GTGTAGGTCG	TTGCTCCAA	GCTGGGCTGT	GTGCACGAAC	CCCCCGTTCA
18481	GCCCGACCGC	TGCGCCTTAT	CCGGTAACTA	TCGTCTTGAG	TCCAACCCGG	TAAGACACGA
18541	CTTATCGCCA	CTGGCAGCAG	CCACTGGTAA	CAGGATTAGC	AGAGCGAGGT	ATGTAGGCGG
18601	TGCTACAGAG	TTCTTGAAGT	GGTGGCCTAA	CTACGGCTAC	ACTAGAAGGA	CAGTATTTGG
18661	TATCTGCGCT	CTGCTGAAGC	CAGTTACCTT	CGGAAAAAGA	GTTGGTAGCT	CTTGATCCGG
18721	CAAAACAAACC	ACCGCTGGTA	GCGGTGGTTT	TTTTGTTTGC	AAGCAGCAGA	TTACGCGCAG
18781	AAAAAAAGGA	TCTCAAGAAG	ATCCTTTGAT	CTTTTCTACG	GGGTCTGACG	CTCAGTGGAA
18841	CGAAAACCTCA	CGTTAAGGGA	TTTTTGGTCAT	GAGATTATCA	AAAAGGATCT	TCACCTAGAT
18901	CCTTTTAAAT	TAAAAATGAA	GTTTTAAATC	AATCTAAAGT	ATATATGAGT	AAACTTGGTC
18961	TGACAGTTAC	CAATGCTTAA	TCAGTGAGGC	ACCTATCTCA	GCGATCTGTC	TATTTGTTTC
19021	ATCCATAGTT	GCCTGACTCC	CCGTCGTGTA	GATAACTACG	ATACGGGAGG	GCTTACCATC
19081	TGGCCCCAGT	GCTGCAATGA	TACCGCGAGA	CCCACGCTCA	CCGGCTCCAG	ATTTATCAGC
19141	AATAAACCAG	CCAGCCGGAA	GGGCCGAGCG	CAGAAAGTGG	CCTGCAACTT	TATCCGCCTC
19201	CATCCAGTCT	ATTAATTGTT	GCCGGGAAGC	TAGAGTAAAGT	AGTTCGCCAG	TTAATAGTTT
19261	GCGCAACGTT	GTTGCCATTG	CTACAGGCAT	CGTGGTGTC	CGCTCGTCGT	TTGGTATGGC
19321	TTCATTAGC	TCCGTTTCCC	AACGATCAAG	GCGAGTTACA	TGATCCCCCA	TGTTGTGCAA
19381	AAAAGCGGTT	AGCTCCTTCG	GTCCTCCGAT	CGTTGTCAGA	AGTAAGTTGG	CCGAGTGTTC
19441	ATCACTCATG	GTTATGGCAG	CACTGCATAA	TTCTCTTACT	GTCATGCCAT	CCGTAAGATG
19501	CTTTTCTGTG	ACTGGTGAGT	ACTCAACCAA	GTCAATCTGA	GAATAGTGTA	TGCGGCGACC
19561	GAGTTGCTCT	TGCCCGGCGT	CAATACGGGA	TAATACCGCG	CCACATAGCA	GAACCTTAAA
19621	AGTGCTCATC	ATTGGAAAAC	GTTCTTCGGG	GCGAAAACCTC	TCAAGGATCT	TACCGCTGTT
19681	GAGATCCAGT	TCGATGTAAC	CCACTCGTGC	ACCCAACTGA	TCTTCAGCAT	CTTTTACTTT
19741	CACCAGCGTT	TCTGGGTGAG	CAAAAACAGG	AAGGCAAAAT	GCCGCAAAAA	AGGGAATAAG
19801	GGCGACACGG	AAATGTTGAA	TACTCATACT	CTTCCTTTTT	CAATATTATT	GAAGCATTTA
19861	TCAGGGTTAT	TGTCTCATGA	GCGGATACAT	ATTTGAATGT	ATTTAGAAAA	ATAAACAAAT
19921	AGGGGTTCCG	CGCACATTTC	CCCGAAAAGT	GCCACCTAAA	TTGTAAGCGT	TAATATTTTG
19981	TTAAAATTCTG	CGTTAAATTT	TTGTTAAATC	AGCTCATTTT	TTAACCAATA	GGCCGAAATC
20041	GGCAAAATCC	CTTATAAATC	AAAAGAATAG	ACCGAGATAG	GGTTGAGTGT	TGTTCCAGTT
20101	TGGAACAAGA	GTCCACTATT	AAAAGAACGTG	GACTCCAACG	TCAAAGGGCG	AAAAACCGTC
20161	TATCAGGGCG	ATGGCCCACT	ACGTGAACCA	TCACCCTAAT	CAAGTTTTTT	GGGGTCGAGG
20221	TGCCGTAAG	CACTAAATCG	GAACCTTAAA	GGGAGCCCCC	GATTTAGAGC	TTGACGGGGA
20281	AAGCCGGCGA	ACGTGGCGAG	AAAGGAAGGG	AAGAAAGCGA	AAGGAGCGGG	CGCTAGGGCG
20341	CTGGCAAGTG	TAGCGGTCAC	GCTGCGCGTA	ACCACCACAC	CCGCCGCGCT	TAATGCGCCG
20401	CTACAGGGCG	CGTCCCATT	GCCATTTCAG	CTGCGCAACT	GTTGGGAAGG	GCGATCGGGT
20461	CGGGCCTCTT	CGCTATTACG	CCAGCTGGCG	AAAGGGGGAT	GTGCTGCAAG	GCGATTAAGT
20521	TGGGTAAACG	CAGGGTTTTT	CCAGTACCGA	CGTTGTAAAA	CGACGGCCAG	TGAGCGCGCG
20581	TAATACGACT	CACTATAGGG	CGAATTGGGT	ACCGGGCCCC	CCC	

**Figure 6**

**Figure 7**

**Figure 8**

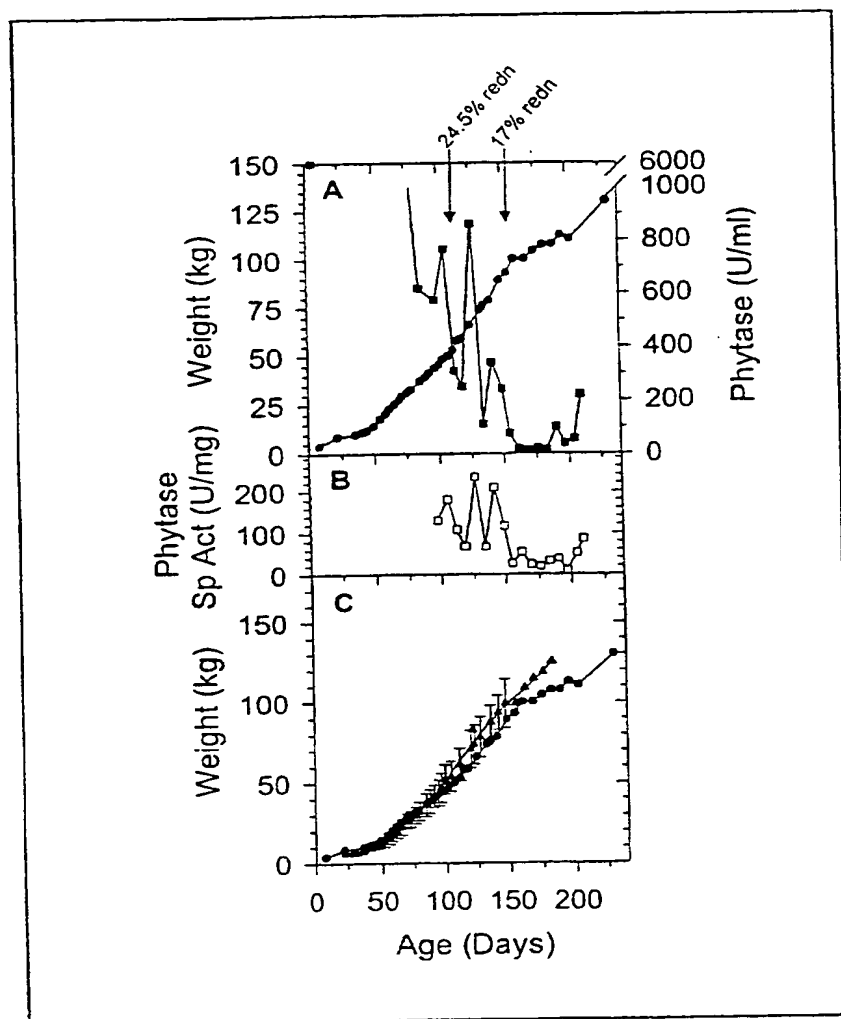


Figure 9

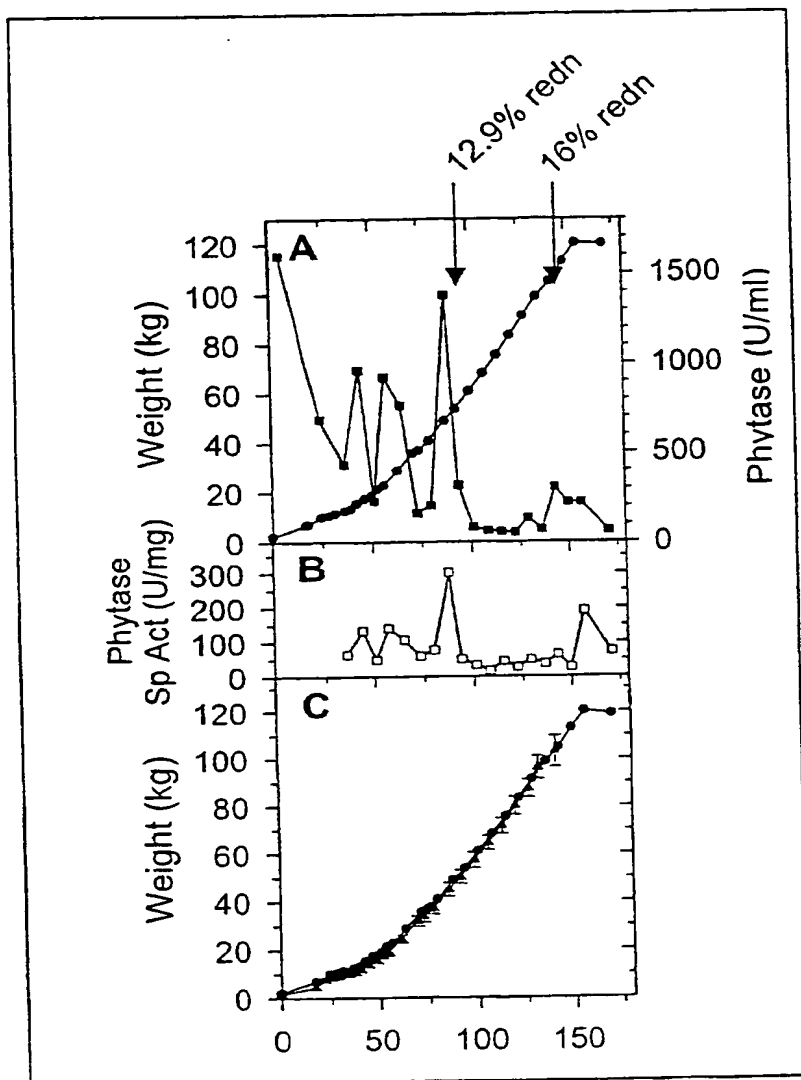


Figure 10

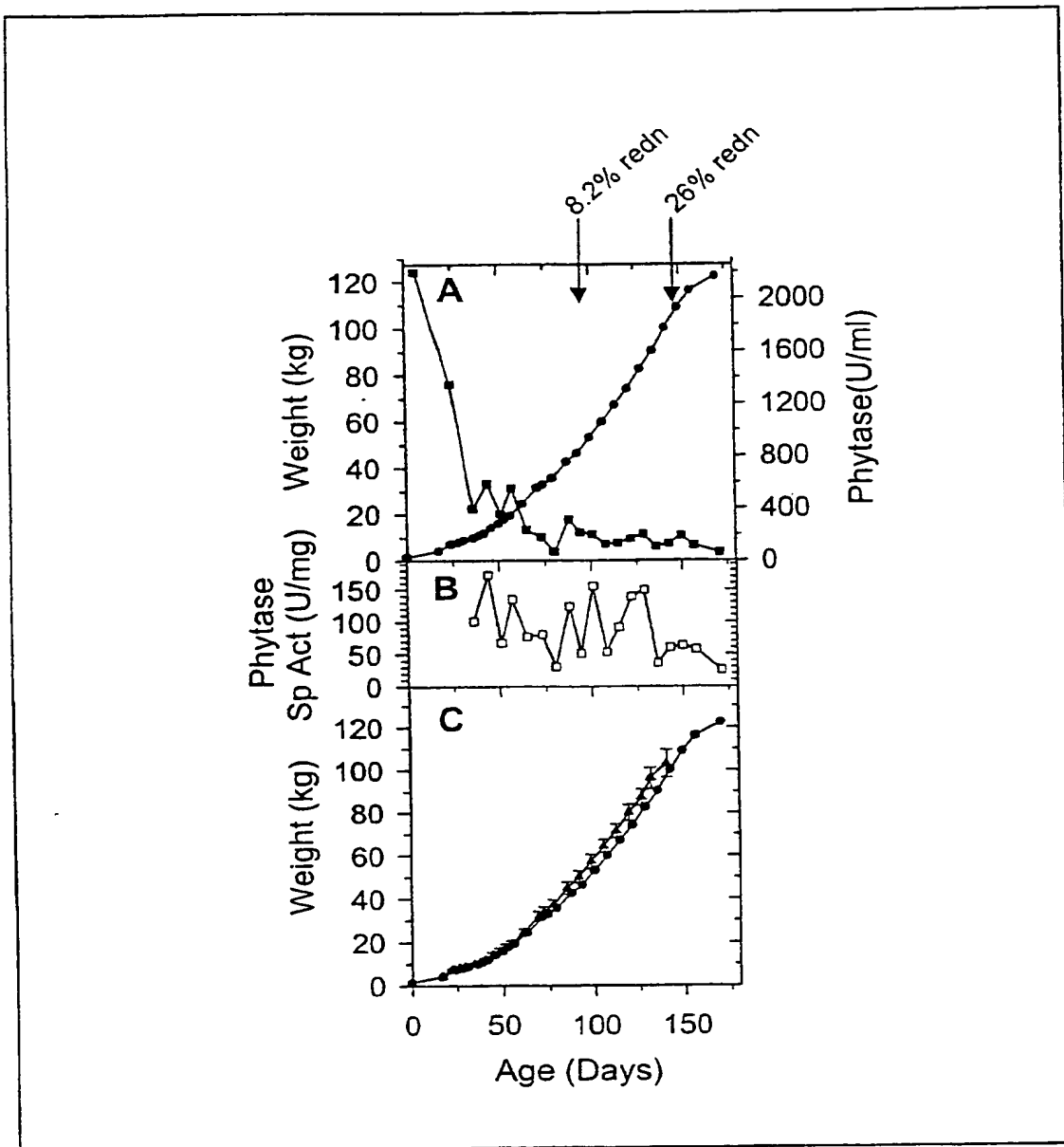
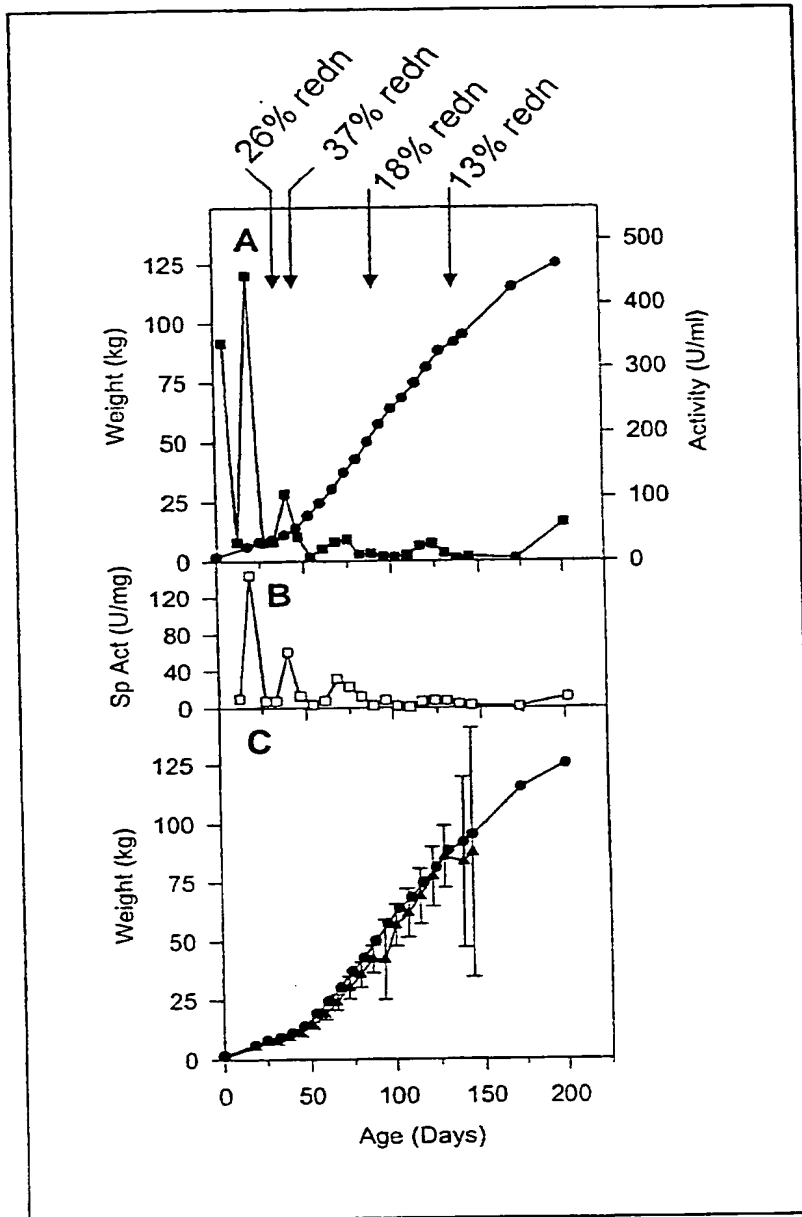


Figure 11

**Figure 12**

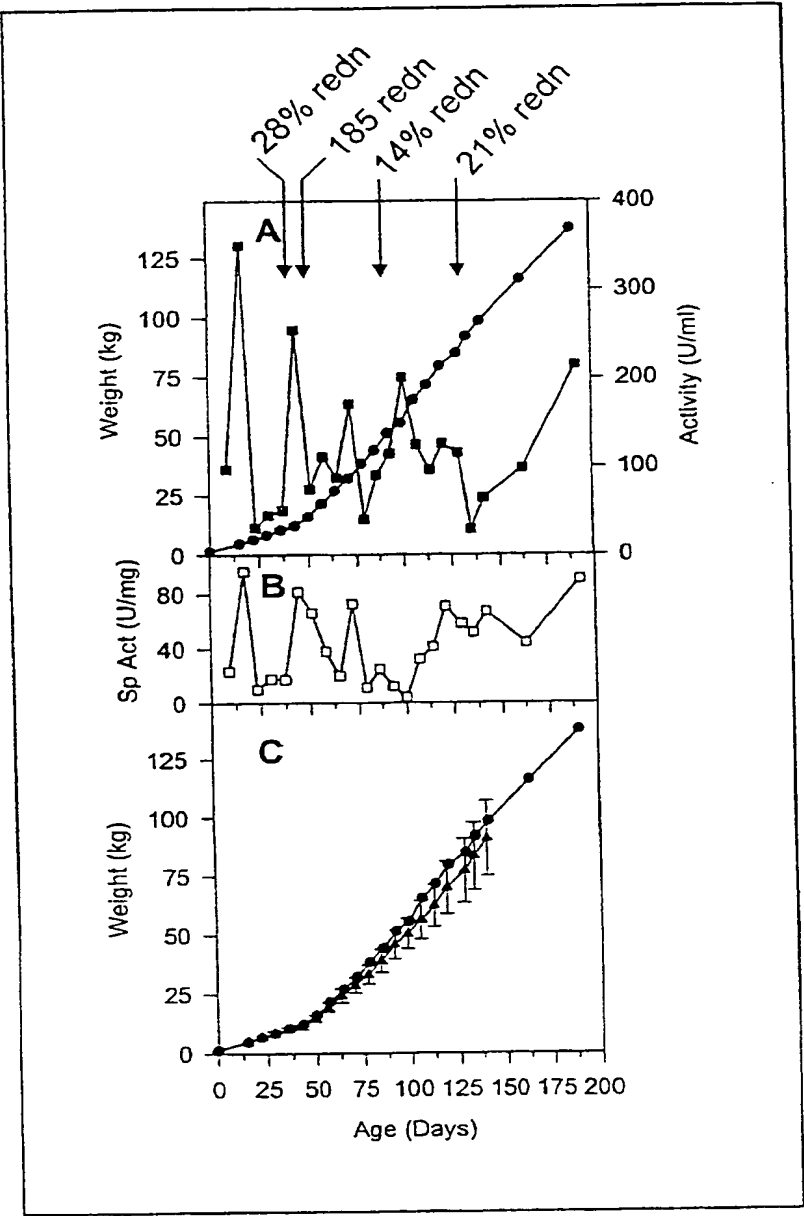


Figure 13

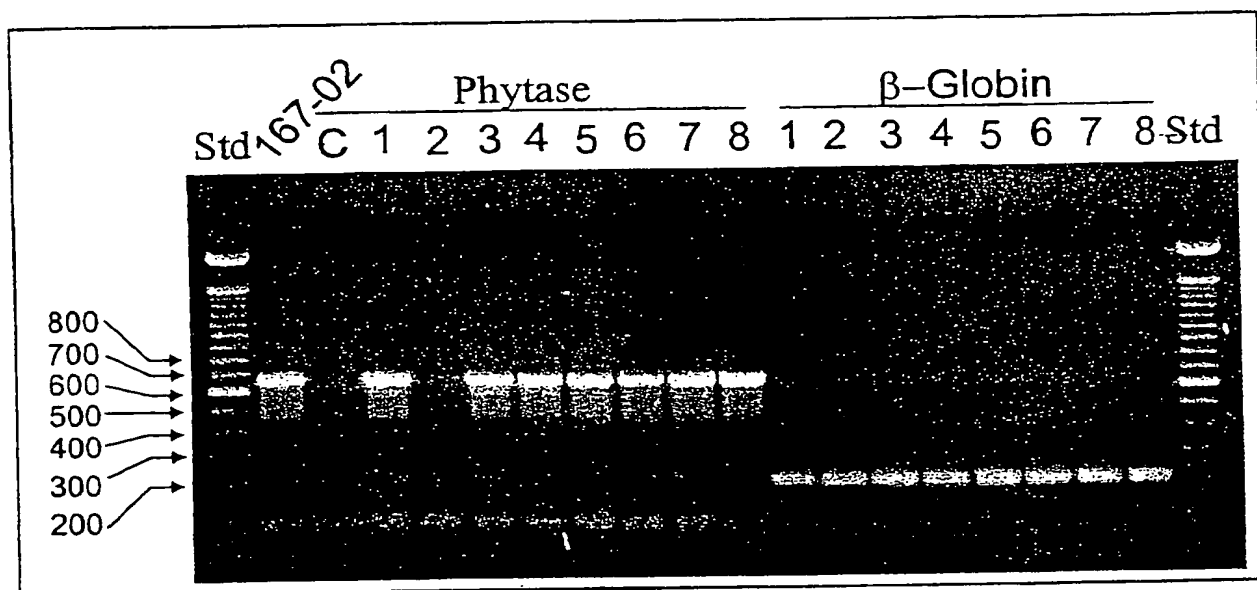


Figure 14

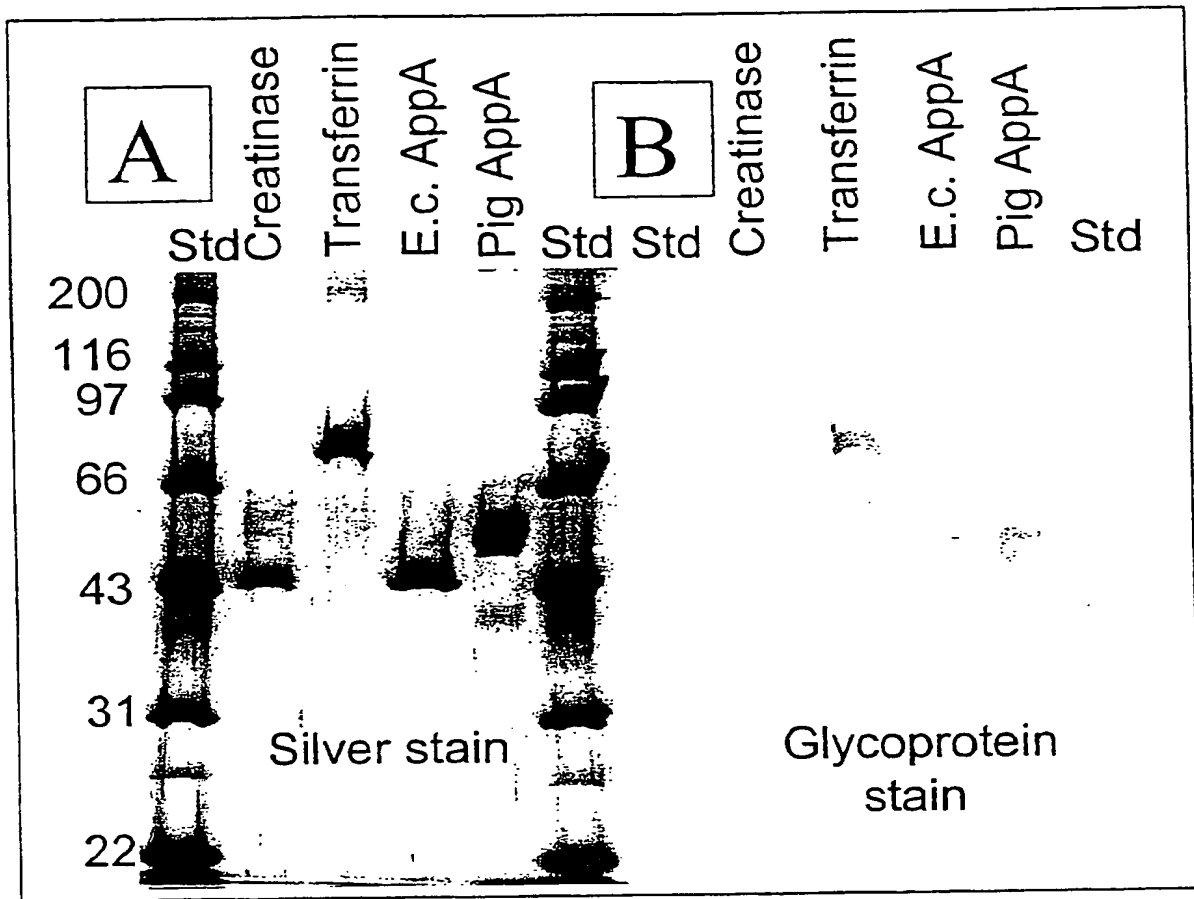


Figure 15

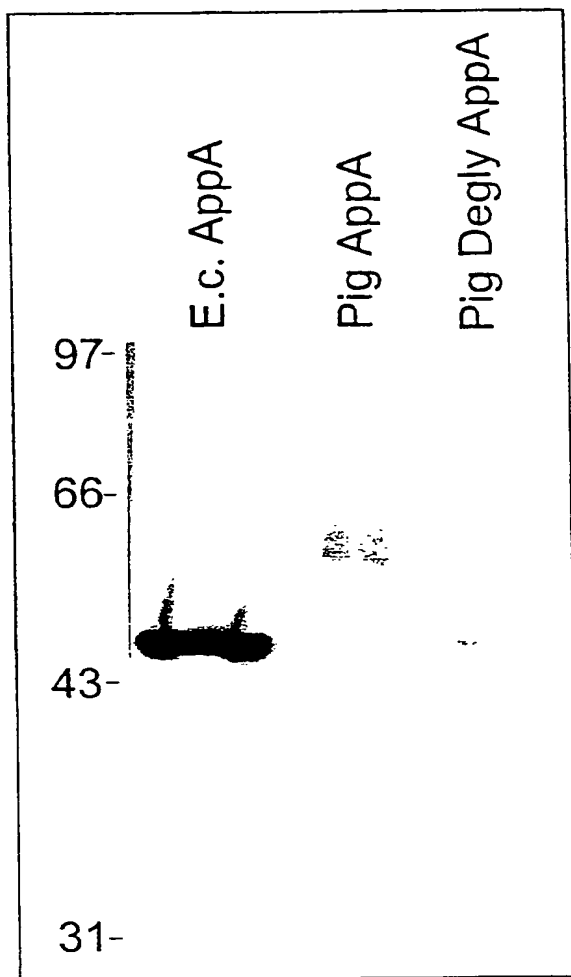


Figure 15B

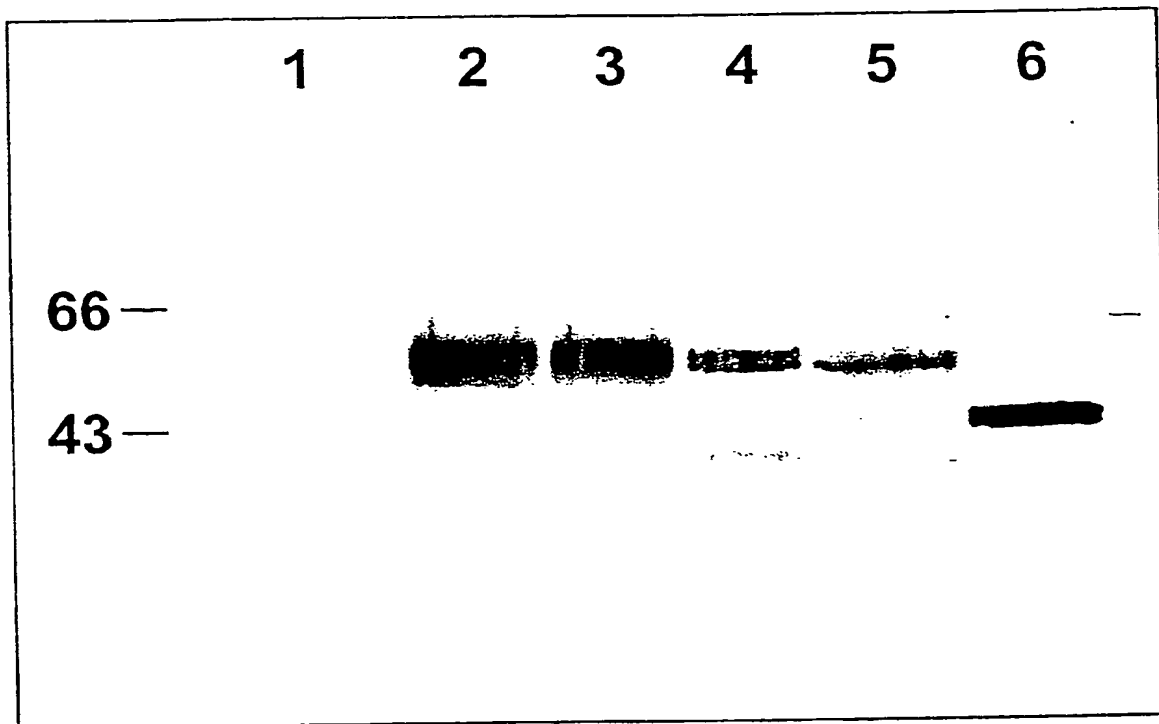


Figure 16

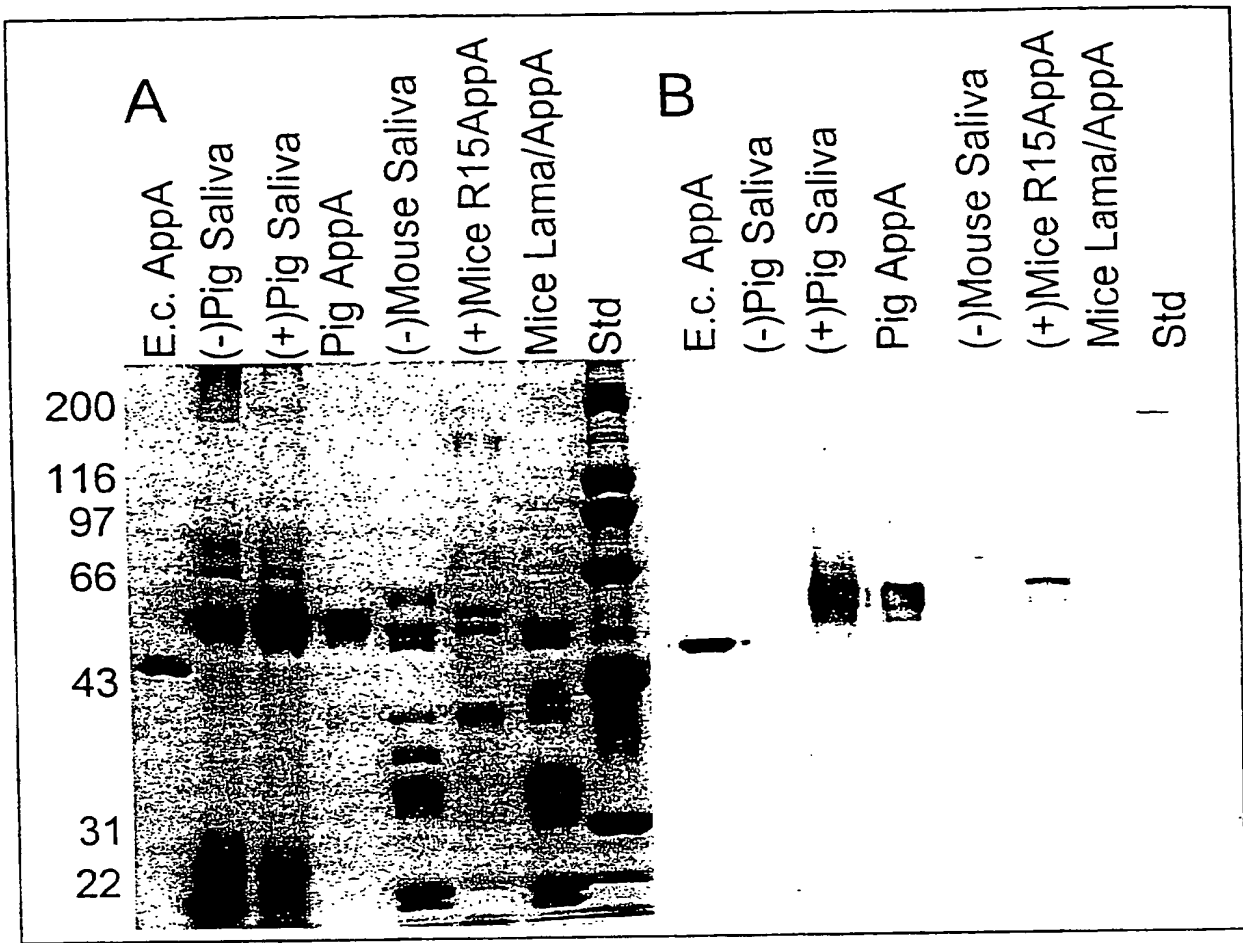


Figure 17

Figure 18: Nucleic acid sequence of the known segment of the R15/appa+intron plasmid, including the vector sequences of pBLCAT3 (SEQ ID NO:2).

LOCUS R15/appa+intron 6708 bp DNA SYN 15-APR-2000
 DEFINITION R15/appa+intron transgene with vector cut 13543 to 4954
 ACCESSION R15/appa+intron
 REFERENCE 1 (bases 1 to 6708))
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 KEYWORDS salivary proline-rich protein, acid glucose-1-phosphatase; appA
 gene; periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

DEFINITION Rat salivary proline-rich protein (RP15) gene.
 ACCESSION M64793 M36414
 VERSION M64793.1 GI:206711
 SOURCE Rat (Sprague-Dawley) liver DNA.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1748)
 AUTHORS Lin, H.H. and Ann, D.K.
 TITLE Molecular characterization of rat multigene family
 encoding
 proline-rich proteins
 JOURNAL Genomics 10, 102-113 (1991)
 MEDLINE 91257817
 FEATURES Location/Qualifiers
 source 1..1748
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 /tissue_lib="cosmid genomic library"
 misc_feature 1802-1810
 /function=" consensus sequence for initiation in
 higher eukaryotes "

FEATURES Location/Qualifiers
 DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
 gene,

ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
 VERSION M58708.1 GI:145283
 SOURCE Escherichia coli DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision;
 Enterobacteriaceae;
 Escherichia.
 REFERENCE 1 (bases 1811..3109)
 AUTHORS Dassa, J., Marck, C. and Boquet, P.L.

Figure 18 (continued):

TITLE The complete nucleotide sequence of the Escherichia coli gene appA reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase

JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)

MEDLINE 90368616

FEATURES Location/Qualifiers

Source 1811..3109
/organism="Escherichia coli"
/db_xref="taxon:562"

sig_peptide 1811.. 1876
/gene="appA"

CDS 1811..3109
/gene="appA"
/standard_name="acid phosphatase/phytase"
/transl_table=11
/product="periplasmic phosphoanhydride phosphohydrolase"
/protein_id="AAA72086.1"
/db_xref="GI:145285"

/translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
TKATQLMQDVTTPDAWPTWPVKLGWLTTPRGGELIAYLGHYQRQLVADGLLAKKGC PQS
GQVAI IADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
YLLQRTPEVARSRATPLLDLIK TALTPHPPQKQAYGVTLP TSVLFIAGHDTNLANLGG
ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT
PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"

mat_peptide 1877 3106
/gene="appA"
/product="periplasmic phosphoanhydride phosphohydrolase"

mutation replace(1817.. 1819,"gcg changed to gcc")
/gene="appA"
/standard_name="A3 mutant"
/note="created by site directed mutagenesis"
/phenotype="silent mutation"

mutation replace(3092..3094," ccg changed to ccc")
/gene="appA"
/standard_name=" P428 mutant"
/note="created by site directed mutagenesis"
/phenotype=" silent mutation "

mutation replace(3095..3097," gcg changed to gct")
/gene="appA"
/standard_name=" A429 mutant"
/note="created by site directed mutagenesis"
/phenotype=" silent mutation "

Figure 18 (continued):

DEFINITION Plasmid pBLCAT3 (bases 3109 to 6708)

ACCESSION X64409
 VERSION X64409.1 GI:58163
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 REFERENCE 1 (bases 3109 to 6708)
 AUTHORS Luckow, B.H.R.
 TITLE Direct Submission
 JOURNAL Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res
 Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG
 REFERENCE 2 (bases 3109 to 6708)
 AUTHORS Luckow, B. and Schutz, G.
 TITLE CAT constructions with multiple unique restriction sites

for

the functional analysis of eukaryotic promoters and
 regulatory elements

JOURNAL Nucleic Acids Res. 15 (13), 5490 (1987)
 MEDLINE 87260024
 COMMENT Promoterless CAT vector for transient transfection
 experiments

with eukaryotic cells. Allows the analysis of foreign
 promoters and enhancers.

FEATURES Location/Qualifiers
 source 3109 to 6116
 /organism="synthetic construct"
 /db_xref="taxon:32630"

SV40 t intron 3197..3810
 /note="SV40 signals"
 polyA_signal 3807..4047
 /note="SV40 signals"
 CDS complement(5244..6104)
 /codon_start=1
 /transl_table=11
 /gene="Amp"
 /product="beta-lactamase"
 /protein_id="CAA45753.1"
 /db_xref="GI:58165"

BASE COUNT 1916 a 1479 c 1515 g 1798 t
 ORIGIN

```

1 GGATCCCCTT TGCTATGTAG TTTTAAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
61 GAGAGTCCTG TTTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
121 CTCTTTGTTT CTAGCATAAC CAAAAGATTT AGTGAATTGA AAACAATGTT CCCTTAGAGT
181 ATAGGTCTAA TAACCCCGAA AATATTACCA TGATACTGAG CATTTGTAAG TATCTCATAG
241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTTAA CATGATTTTC ATTAATCAGG
301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGCACA
361 TATTTCACTA AACTAGGTTT ATCTATTTTG TTGCTTTCTC TAACATCTCT GCAATGAAGC
421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAT TTTCTACATA
481 TATCCTGGTT AGAGAGTGCT TAAAATAAGT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT
601 TGGGAAGAAA CCATTTGGTG AACAATATTT CAAATAAAAA TAGACAAACA TAGTTAATTG
661 TAAAAACATAT GTTTGACCAG CCCTTCTTTT CAATAGGCTT AATGTGAATA AAATGTTAAA
721 GATTCTCTTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT
841 CACAAATTAA TCAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT

```

Figure 18 (continued):

```

901 TAAGATAAAG GTAACGTGTAT ACATTTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
961 TTCAGCTCTA TAATTCTTGC CTTAAACAAC TTAAATAGAA TGATTTAAAA TATGGAGCTG
1021 TCCATGGACC TTTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCCCTT
1081 GATACTAACA CAGGTAAATC CCACACGTGT TTTGAGACTA CATTTGCTGG GATTTTATTG
1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTTGCTGTT ATTCGGTTAC TTCAAGAGAA
1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATTT
1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
1321 GTGTTTAAGC TGTACTATTG ATCAAAGAAA TTTATTACCT TCAGTTTCAA TGGAAATAAT
1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAATCCTA TCCAAATGCA CAGTGATACA
1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTTA TCAATCAATT
1501 GTATGTATCA ATATATGGGC TATTTTCTTA CACATGATTT TATTCAAAT TACTCTAATC
1561 ATTGTTGAAC CATTTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGGC
1621 AAAAGTCCCA GTGTGGAGTA AAGGATGCAA GATTTCCCTGC TCTGTTAAGT ATAAAATAAT
1681 AGTATGAATT CAAAGGTGCC ATTCTTCTGC TTCTAGTTAT AAAGGCAGTG CTTGCTTCTT
1741 CCAGCACAGA TCTGGATCTC GAGGAGCTTG GCGAGATTTT CAGGAGCTAA GGAAGCTAAA
1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTTTATCT CTTCTGATT CGTAAACCCC
1861 GCAATCTGCA TTCGCTCAGA GTGAGCCGGA GCTGAAGCTG GAAAGTGTGG TGATTGTCAG
1921 TCGTCATGGT GTGCGTGCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCGAGA
1981 CGCATGGCCA ACCTGGCCGG TAAAACTGGG TTGGCTGACA CCGCGCGGTG GTGAGCTAAT
2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCGAAAAA
2101 GGGCTGCCCG CAGTCTGGTC AGGTCGCGAT TATTGCTGAT GTCGACGAGC GTACCCGTAA
2161 AACAGGCGAA GCCTTCGCCG CCGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCCA
2221 GGCAGATACG TCCAGTCCCG ATCCGTTATT TAATCCTCTA AAAACTGGCG TTTGCCAACT
2281 GGATAACGCG AACGTGACTG ACGCGATCCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
2341 TACCGGGCAT CGGCAAACGG CGTTTCGCGA ACTGGAACGG GTGCTTAATT TTCCGCAATC
2401 AAACCTTGTC CTTAAACGTG AGAAACAGGA CGAAAGCTGT TCATTAACGC AGGCATTACC
2461 ATCGGAACCT AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC
2521 AATGCTGACG GAGATATTTT TCCTGCAACA AGCACAGGGA ATGCCGAGC CGGGGTGGGG
2581 AAGGATCACC GATTACACAC AGTGGAACAC CTTGCTAAGT TTGCATAACG CGCAATTTTA
2641 TTTGCTACAA CGCACGCCAG AGGTTGCCCG CAGCCGCGCC ACCCGTTAT TAGATTTGAT
2701 CAAGACAGCG TTGACGCCCC ATCCACCGCA AAAACAGGCG TATGGTGTGA CATTACCCAC
2761 TTCAGTGCTG TTTATCGCCG GACACGATAC TAATCTGGCA AATCTCGGCG GCGCACTGGA
2821 GCTCAACTGG ACGCTTCCCG GTCAGCCGGA TAACACGCCG CCAGGTGGTG AACTGGTGT
2881 TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGATT CAGGTTTCGC TGGTCTTCCA
2941 GACTTTACAG CAGATGCGTG ATAAAACGCC GCTGTCATTA AATACGCCG CCGGAGAGGT
3001 GAAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGCAG GGCATGTGTT CGTTGGCAGG
3061 TTTTACGCAA ATCGTGAATG AAGCAGCAT ACCCGCTTGC AGTTTGTAAG GTATAAGGCA
3121 GTTATTGGTG CCCTTAAACG CCTGGTGCTA CGCCTGAATA AGTGATAATA AGCGGATGAA
3181 TGGCAGAAAT TCGCCGGATC TTTGTGAAGG AACCTTACTT CTGTGGTGTG ACATAATTGG
3241 ACAAACCTACC TACAGAGATT TAAAGCTCTA AGGTAAATAT AAAATTTTTA AGTGATATAAT
3301 GTGTAAACT ACTGATTCTA ATTGTTGTG TATTTTAGAT TCCAACCTAT GGAAGTATG
3361 AATGGGAGCA GTGGTGGAAT GCCTTTAATG AGGAAAACCT GTTTTGCTCA GAAGAAATGC
3421 CATCTAGTGA TGATGAGGCT ACTGCTGACT CTCACATTC TACTCCTCCA AAAAAGAAGA
3481 GAAAGGTAGA AGACCCCAAG GACTTTCCTT CAGAATTGCT AAGTTTTTGT AGTCATGCTG
3541 TGTTTAGTAA TAGAACTCTT GCTTGCTTTG CTATTTACAC CACAAAGGAA AAAGCTGCAC
3601 TGCTATACAA GAAAATTATG GAAAAATATT CTGTAACCTT TATAAGTAGG CATAACAGTT
3661 ATAATCATAA CATACTGTTT TTTCTTACTC CACACAGGCA TAGAGTGTCT GCTATTAATA
3721 ACTATGCTCA AAAATTGTGT ACCTTTAGCT TTTTAATTTG TAAAGGGGTT AATAAGGAAT
3781 ATTTGATGTA TAGTGCCTTG ACTAGAGATC ATAATCAGCC ATACCACATT TGTAGAGGTT
3841 TTAAGTGTG TAAAAACCT CCCACACCTC CCCCTGAACC TGAAACATAA AATGAATGCA
3901 ATTGTTGTTG TTAAGTGTG TATTGCAGCT TATAATGGTT ACAAATAAAG CAATAGCATC
3961 ACAAATTTCA CAAATAAAGC ATTTTTCCTA CTGCATTCTA GTTGTGGTTT GTCCAAACTC
4021 ATCAATGTAT CTTATCATGT CTGGATCGAT CCCCAGGTAC CGAGCTCGAA TTCGTAATCA
4081 TGGTCATAGC TGTTTCCTGT GTGAAATTGT TATCCGCTCA CAATTCCACA CAACATACGA
4141 GCCGGAAGCA TAAAGTGTA AGCCTGGGTG GCCTAATGAG TGAGCTAAGT CACATTAATT
4201 GCGTTGCGCT CACTGCCCGC TTTCCAGTCG GGAAACCTGT CGTGCCAGCT GCATTAATGA
4261 ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTTCGCG TTCTCGCTC
4321 ACTGACTCGC TGCGCTCGGT CGTTCGGCTG CGGCGAGCGG TATCAGCTCA CTCAAAGGCG

```

Figure 18 (continued):

```

4381 GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC
4441 CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC
4501 CCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA
4561 CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGTCCCCTCG TCGCTCTCC GTTCCGACC
4621 CTGCCGCTTA CCGGATACCT GTCCGCTTT CTCCCTTCGG GAAGCGTGGC GCTTCTCAA
4681 TGCTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTT GCTCCAAGCT GGGCTGTGTG
4741 CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAAGTATCG TCTTGAGTCC
4801 AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA
4861 GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT
4921 AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT
4981 GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTGTGCAAG
5041 CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG
5101 TCTGACGCTC AGTGAACGA AAATCAGCT TAAGGGATTT TGGTCATGAG ATTATCAAAA
5161 AGGATCTTCA CCTAGATCCT TTAAATTA AAATGAAGTT TAAATCAAT CTAAAGTATA
5221 TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG
5281 ATCTGTCTAT TTCGTTTCAT CATATTGCCC TGACTCCCCG TCGTGTAGAT AACTACGATA
5341 CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG
5401 GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT
5461 GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAGCTAG AGTAAGTAGT
5521 TCGCCAGTTA ATAGTTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC
5581 TCGTCGTTTG GTATGGCTTC ATTCAGCTCC GGTTCCTAAC GATCAAGGCG AGTTACATGA
5641 TCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTGAGAGT
5701 AAGTTGGCCG CAGTGTATAT ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC
5761 ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA
5821 TAGTGTATGC GCGGACCGAG TTGCTCTTGC CCGGCGTCAA TACGGGATAA TACCGCGCCA
5881 CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTGCGGGCG AAAACTCTCA
5941 AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT
6001 TCAGCATCTT TTACTTTCAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC
6061 GCAAAAAAGG GAATAAGGGC GACACGGAAG TGTGAATAC TCATACTCTT CCTTTTCAA
6121 TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT
6181 TAGAAAAATA AACAAATAGG GGTTCGCGC ACATTTCCCC GAAAAGTGCC ACCTGACGTC
6241 TAAGAAACCA TTATTATCAT GACATTAACC TATAAAAAATA GGCGTATCAC GAGGCCCTTT
6301 CGTCTCGCGC GTTTCGGTGA TGACGGTGAA AACCTCTGAC ACATGCAGCT CCCGGAGACG
6361 GTCACAGCTT GTCTGTAAGC GGATGCCGGG AGCAGACAAG CCCGTCAGGG CGCGTCAGCG
6421 GGTGTTGGCG GGTGTCGGGG CTGGCTTAAC TATGCGGCAT CAGAGCAGAT TGTACTGAGA
6481 GTGCACCATA TCGGTGTGA AATACCGCAC AGATGCGTAA GGAGAAAAATA CCGCATCAGG
6541 CGCCATTTCG CATTACGGCT CGCAACTGT TGGGAAGGGC GATCGGTGCG GGCCTCTTCG
6601 CTATTACGCC AGCTGGCGAA AGGGGGATGT GCTGCAAGGC GATTAAGTTG GGTAACGCCA
6661 GGGTTTTCCC AGTCACGACG TTGTAAAACG ACGGCCAGTG CCAAGCTT

```

//

Figure 19: Nucleic acid sequence of the known segment of the R15/appa-intron transgene used for the generation of transgenic mice (SEQ ID NO: 3).

LOCUS R15/appa 4060 bp DNA SYN 15-APR-2000
 DEFINITION R15/appa transgene without vector
 ACCESSION R15/appa
 REFERENCE 1 (bases 1 to 4060)
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 KEYWORDS salivary proline-rich protein, acid glucose-1-phosphatase; appA
 gene; periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

DEFINITION Rat salivary proline-rich protein (RP15) gene.
 ACCESSION M64793 M36414
 VERSION M64793.1 GI:206711
 SOURCE Rat (Sprague-Dawley) liver DNA.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 1748)
 AUTHORS Lin, H.H. and Ann, D.K.
 TITLE Molecular characterization of rat multigene family
 encoding

proline-rich proteins
 JOURNAL Genomics 10, 102-113 (1991)
 MEDLINE 91257817
 FEATURES Location/Qualifiers
 source 1..1748
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 /tissue_lib="cosmid genomic library"
 misc_feature 1802-1810
 /function=" consensus sequence for initiation in
 higher eukaryotes "

FEATURES Location/Qualifiers
 DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
 gene,

ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
 VERSION M58708.1 GI:145283
 SOURCE Escherichia coli DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision;
 Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (bases 1811..3109)
 AUTHORS Dassa, J., Marck, C. and Boquet, P.L.

Figure 19 (continued):

TITLE The complete nucleotide sequence of the Escherichia coli
gene appA reveals significant homology between pH 2.5
acid phosphatase and glucose-1-phosphatase

JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)

MEDLINE 90368616

FEATURES Location/Qualifiers

Source 1811..3109
/organism="Escherichia coli"
/db_xref="taxon:562"

sig_peptide 1811..1876
/gene="appA"

CDS 1811..3109
/gene="appA"
/standard_name="acid phosphatase/phytase"
/transl_table=11
/product="periplasmic phosphoanhydride
phosphohydrolase"
/protein_id="AAA72086.1"
/db_xref="GI:145285"

/translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
TKATQLMQDVTTPDAWPTWPVKLGWLTTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS
GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
YLLQRTPEVARSRATPLLDLIKALTTPHPPQKQAYGVTLPSTSVLFIAGHDTNLANLGG
ALELNWTLPGQPDNTPPGGELVFERWRRRLSDNSQWIVSLVFQTLQQMRDKTPLSLNT
PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"

mat_peptide 1877 3106
/gene="appA"
/product="periplasmic phosphoanhydride
phosphohydrolase"

mutation replace(1817..1819,"gcg changed to gcc")
/gene="appA"
/standard_name="A3 mutant"
/note="created by site directed mutagenesis"
/phenotype="silent mutation"

mutation replace(3092..3094,"ccg changed to ccc")
/gene="appA"
/standard_name="P428 mutant"
/note="created by site directed mutagenesis"
/phenotype="silent mutation"

mutation replace(3095..3097,"gcg changed to gct")
/gene="appA"
/standard_name="A429 mutant"
/note="created by site directed mutagenesis"
/phenotype="silent mutation"

Figure 19 (continued):

SV40 t intron 3197..3810
/note="SV40 signals"
polyA_signal 3807..4047
/note="SV40 signals"

BASE COUNT 1257 a 814 c 843 g 1146 t
ORIGIN

```

1  GGATCCCCTT TGCTATGTAG TTTTAAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
61 GAGAGTCCTG TTTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
121 CTCTTTGTTT CTAGCATAAC CAAAAGATTT AGTGAATTGA AAACAATGTT CCCTTAGAGT
181 ATAGGTCTAA TAACCCCGAA AATATTACCA TGATACTGAG CATTTGTAAG TATCTCATAG
241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTTAA CATGATTTTC ATTAATCAGG
301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGACAC
361 TATTTCACTA AACTAGGTTT ATCTATTTTG TTGCTTTCTC TAACATCTCT GCAATGAAGC
421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAT TTTCTACATA
481 TATCCTGGTT AGAGAGTGCT TAAAATAAGT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT
601 TGGGAAGAAA CCATTGGTG AACAATATTT CAAATAAAAA TAGACAAACA TAGTTAATTG
661 TAAACATAT GTTTGACCAG CCCTTCTTTT CAATAGGCTT AATGTGAATA AAATGTTAAA
721 GATTCTCTTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAAGAGA ATCTTGAGAG AATGTGTTGT
841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT
901 TAAGATAAAG GTAACGTAT ACATTTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
961 TTCAGCTCTA TAATCTTGC CTTAAACAAC TTAAATAGAA TGATTTAAAA TATGGAGCTG
1021 TCCATGGACC TTTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCCTT
1081 GATACTAACA CAGGTAAATC CCACACGTGT TTTGAGACTA CATTTGCTGG GATTTTATTG
1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTTGCTGTT ATTCGGTTAC TTCAAGAGAA
1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATTT
1261 GTAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
1321 GTGTTTAAGC TGTACTATTG ATCAAAGAAA TTTATTACCT TCAGTTTCAA TGGAAATAAT
1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAATCCTA TCCAAATGCA CAGTGATACA
1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTTA TCAATCAATT
1501 GTATGTATCA ATATATGGGC TATTTTCTTA CACATGATTT TATTCAAAT TACTCTAATC
1561 ATTGTTGAAC CATTTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGGC
1621 AAAAGTCCCA GTGTGGAGTA AAGGATGCAA GATTTCTGTC TCTGTTAAGT ATAAAATAAT
1681 AGTATGAATT CAAAGGTGCC ATTCTTCTGC TTCTAGTTAT AAAGGCAGTG CTGCTTCTT
1741 CCAGCACAGA TCTGGATCTC GAGGAGCTTG GCGAGATTTT CAGGAGCTAA GGAAGCTAAA
1801 AGCCGCCACC ATGAAAAGCCA TCTTAATCCC ATTTTATCT CTCTGATTC CGTTAACCCC
1861 GCAATCTGCA TTCGCTCAGA GTGAGCCGGA GCTGAAGCTG GAAAGTGTGG TGATTGTCAG
1921 TCGTCATGGT GTGCGTGCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCAGA
1981 CGCATGGCCA ACCTGGCCGG TAAAAGTGGG TTGGCTGACA CCGCGCGGTG GTGAGCTAAT
2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCGAAAAA
2101 GGGCTGCCCG CAGTCTGGTC AGGTGCGCAT TATTGCTGAT GTCGACGAGC GTACCCGTAA
2161 AACAGGCGAA GCCTTCGCCG CCGGCTGGC ACCTGACTGT GCAATAACCG TACATACCCA
2221 GGCAGATACG TCCAGTCCCG ATCCGTTATT TAATCCTCTA AAAAGTGGCG TTTGCCAACT
2281 GGATAACGCG AACGTGACTG AACGATCCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
2341 TACCGGGCAT CGGCAAACGG CGTTTCGCGA ACTGGAACGG GTGCTTAATT TTCCGCAATC
2401 AAACCTGTGC CTTAAACGTG AGAAACAGGA CGAAAGCTGT TCATTAACGC AGGCATTACC
2461 ATCGGAACTC AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC
2521 AATGCTGACG GAGATATTTT TCCTGCAACA AGCACAGGGA ATGCCGAGC CGGGGTGGGG
2581 AAGGATCACC GATTACACC AGTGAACAC CTTGCTAAGT TTGCATAACG CGCAATTTTA
2641 TTTGCTACAA CGCACGCCAG AGGTTGCCCG CAGCCGCGCC ACCCGGTTAT TAGATTTGAT
2701 CAAGACAGCG TTGACGCCCC ATCCACCGCA AAAACAGGCG TATGGTGTGA CATTACCCAC
2761 TTCAGTGCTG TTTATCGCCG GACACGATAC TAATCTGGCA AATCTCGGCG GCGCACTGGA
2821 GCTCAACTGG ACGCTTCCCG GTCAGCCGGA TAACACGCCG CCAGGTGGTG AACTGGTGT
2881 TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGGATT CAGGTTTCGC TGCTCTTCCA

```

Figure 19 (continued):

2941	GACTTTACAG	CAGATGCGTG	ATAAAACGCC	GCTGTCATTA	AATACGCCGC	CCGGAGAGGT
3001	GAAACTGACC	CTGGCAGGAT	GTGAAGAGCG	AAATGCGCAG	GGCATGTGTT	CGTTGGCAGG
3061	TTTTACGCAA	ATCGTGAATG	AAGCACGCAT	ACCCGCTTGC	AGTTTGTAAG	GTATAAGGCA
3121	GTTATTGGTG	CCCTTAAACG	CCTGGTGCTA	CGCCTGAATA	AGTGATAATA	AGCGGATGAA
3181	TGGCAGAAAT	TCGCCGGATC	TTTGTGAAGG	AACCTTACTT	CTGTGGTGTG	ACATAATTGG
3241	ACAAACTACC	TACAGAGATT	TAAAGCTCTA	AGGTAAATAT	AAAATTTTTA	AGTGATAAAT
3301	GTGTTAAACT	ACTGATTCTA	ATTGTTTGTG	TATTTTAGAT	TCCAACCTAT	GGAAGTGATG
3361	AATGGGAGCA	GTGGTGGAAT	GCCTTTAATG	AGGAAAACCT	GTTTTGCTCA	GAAGAAATGC
3421	CATCTAGTGA	TGATGAGGCT	ACTGCTGACT	CTCAACATTC	TACTCCTCCA	AAAAAGAAGA
3481	GAAAGGTAGA	AGACCCCAAG	GACTTTCCTT	CAGAATTGCT	AAGTTTTTTG	AGTCATGCTG
3541	TGTTTAGTAA	TAGAACTCTT	GCTTGCTTTG	CTATTTACAC	CACAAAGGAA	AAAGCTGCAC
3601	TGCTATACAA	GAAAATTATG	GAAAAATATT	CTGTAACCTT	TATAAGTAGG	CATAACAGTT
3661	ATAATCATAA	CATACTGTTT	TTTCTTACTC	CACACAGGCA	TAGAGTGTCT	GCTATTAATA
3721	ACTATGCTCA	AAAATTGTGT	ACCTTTAGCT	TTTTAATTTG	TAAAGGGGTT	AATAAGGAAT
3781	ATTTGATGTA	TAGTGCCTTG	ACTAGAGATC	ATAATCAGCC	ATACCACATT	TGTAGAGGTT
3841	TTACTTGCTT	TAAAAAACCT	CCCACACCTC	CCCCTGAACC	TGAAACATAA	AATGAATGCA
3901	ATTGTTGTTG	TTAACTTGTT	TATTGCAGCT	TATAATGGTT	ACAAATAAAG	CAATAGCATC
3961	ACAAATTTCA	CAAATAAAGC	ATTTTTTTCA	CTGCATTCTA	GTTGTGGTTT	GTCCAAACTC
4021	ATCAATGTAT	CTTATCATGT	CTGGATCGAT	CCCCGGGTAC		

//

Figure 20: Nucleic acid sequence of the known segment of the R15/appa plasmid (including the vector sequences of pBLCAT3 (SEQ ID NO:4)).

LOCUS R15/appa 6116 bp DNA SYN 15-APR-2000
 DEFINITION R15/appa transgene with vector
 ACCESSION R15/appa
 REFERENCE 1 (bases 1 to 6116)
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 KEYWORDS salivary proline-rich protein, acid glucose-1-phosphatase; appA
 gene; periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

DEFINITION Rat salivary proline-rich protein (RP15) gene.
 ACCESSION M64793 M36414
 VERSION M64793.1 GI:206711
 SOURCE Rat (Sprague-Dawley) liver DNA.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1748)
 AUTHORS Lin, H.H. and Ann, D.K.
 TITLE Molecular characterization of rat multigene family
 encoding
 proline-rich proteins
 JOURNAL Genomics 10, 102-113 (1991)
 MEDLINE 91257817
 FEATURES Location/Qualifiers
 source 1..1748
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 /tissue_lib="cosmid genomic library"
 misc_feature 1802-1810
 /function="consensus sequence for initiation in
 higher eukaryotes "

FEATURES Location/Qualifiers
 DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
 gene,
 ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
 VERSION M58708.1 GI:145283
 SOURCE Escherichia coli DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (bases 1811..3109)
 AUTHORS Dassa, J., Marck, C. and Boquet, P.L.
 TITLE The complete nucleotide sequence of the Escherichia coli gene appA
 reveals significant homology between pH 2.5 acid phosphatase
 and glucose-1-phosphatase
 JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)

Figure 20 (continued):

MEDLINE 90368616

FEATURES

Source	Location/Qualifiers
	1811..3109
	/organism="Escherichia coli"
	/db_xref="taxon:562"
sig_peptide	1811..1876
/gene="appA"	
CDS	1811..3109
	/gene="appA"
	/standard_name="acid phosphatase/phytase"
	/transl_table=11
	/product="periplasmic phosphoanhydride phosphohydrolase"
	/protein_id="AAA72086.1"
	/db_xref="GI:145285"

/translation="MKAILIPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
TKATQLMQDVTDPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQLVADGLLAKKGCPQS
GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRI TDSHQWNTLLSLHNAQF
YLLQRTPEVARSRATPLLDLIKALTTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG
ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQOMRDKTPLSLNT
PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"

mat_peptide	1877 3106
	/gene="appA"
	/product="periplasmic phosphoanhydride phosphohydrolase"

mutation	replace(1817..1819,"gcg changed to gcc")
	/gene="appA"
	/standard_name="A3 mutant"
	/note="created by site directed mutagenesis"
	/phenotype="silent mutation"
mutation	replace(3092..3094," ccg changed to ccc")
	/gene="appA"
	/standard_name=" P428 mutant"
	/note="created by site directed mutagenesis"
	/phenotype=" silent mutation "
mutation	replace(3095..3097," gcg changed to gct")
	/gene="appA"
	/standard_name=" A429 mutant"
	/note="created by site directed mutagenesis"
	/phenotype=" silent mutation "

DEFINITION Plasmid pBLCAT3 (bases 3109 to 6116)

ACCESSION	X64409
VERSION	X64409.1 GI:58163
SOURCE	synthetic construct.
ORGANISM	synthetic construct artificial sequence.
REFERENCE	1 (bases 3109 to 6116)
AUTHORS	Luckow, B.H.R.
TITLE	Direct Submission
JOURNAL	Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG

Figure 20 (continued):

REFERENCE 2 (bases 3109 to 6116)
 AUTHORS Luckow, B. and Schutz, G.
 TITLE CAT constructions with multiple unique restriction sites
 for the functional analysis of eukaryotic promoters and
 regulatory elements
 JOURNAL Nucleic Acids Res. 15 (13), 5490 (1987)
 MEDLINE 87260024
 COMMENT Promoterless CAT vector for transient transfection
 experiments with eukaryotic cells. Allows the analysis of foreign
 promoters and enhancers.

FEATURES Location/Qualifiers
 source 3109 to 6116
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 polyA_signal 3262..3457
 /note="SV40 signals"
 CDS complement(4654..5514)
 /codon_start=1
 /transl_table=11
 /gene="Amp"
 /product="beta-lactamase"
 /protein_id="CAA45753.1"
 /db_xref="GI:58165"

BASE COUNT 1724 a 1386 c 1407 g 1599 t
 ORIGIN

```

1  GGATCCCCTT TGCTATGTAG TTTTAAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
61 GAGAGTCCTG TTTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
121 CTCTTTGTTT CTAGCATAAC CAAAAGATTT AGTGAATTGA AAACAATGTT CCCTTAGAGT
181 ATAGGTCTAA TAACCCCGAA AATATTACCA TGATACTGAG CATTTGTAAG TATCTCATAG
241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTTAA CATGATTTTC ATTAATCAGG
301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGCACA
361 TATTTCACTA AACTAGGTTT ATCTATTTTG TTGCTTCTC TAACATCTCT GCAATGAAGC
421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAT TTTCTACATA
481 TATCCTGGTT AGAGAGTGCT TAAAATAAGT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT
601 TGGGAAGAAA CCATTTGGTG AACAATATTT CAAATAAAAA TAGACAAACA TAGTTAATTG
661 TAAACATAT GTTTGACCAG CCCTTCTTTT CAATAGGCTT AATGTGAATA AAATGTTAAA
721 GATTCTCTTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAAGAGA ATCTTGAGAG AATGTGTTGT
841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT
901 TAAGATAAAG GTAACGTAT ACATTTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
961 TTCAGCTCTA TAATCTTGC CTTAAACAAC TTAAATAGAA TGATTTAAAA TATGGAGCTG
1021 TCCATGGACC TTTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCCTT
1081 GATACTAACA CAGGTAAATC CCACACGTGT TTTGAGACTA CATTGCTGG GATTTTATTG
1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTTGCTGTT ATTCGGTTAC TTCAAGAGAA
1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATTT
1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
1321 GTGTTTAAAG TGTACTATTG ATCAAAGAAA TTTATTACCT TCAGTTTCAA TGGAAATAAT
1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAATCCTA TCCAAATGCA CAGTGATACA
1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTTA TCAATCAATT
1501 GTATGTATCA ATATATGGGC TATTTTCTTA CACATGATTT TATTCAAATT TACTCTAATC
1561 ATTGTTGAAC CATTTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGGC
1621 AAAAGTCCCA GTGTGGAGTA AAGGATGCAA GATTTCTGTC TCTGTTAAGT ATAAAATAAT

```

Figure 20 (continued):

```

1681 AGTATGAATT CAAAGGTGCC ATTCTTCTGC TTCTAGTTAT AAAGGCAGTG CTTGCTTCTT
1741 CCAGCACAGA TCTGGATCTC GAGGAGCTTG GCGAGATTTT CAGGAGCTAA GGAAGCTAAA
1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTTTATCT CTTCTGATTC CGTTAACCCC
1861 GCAATCTGCA TTCGCTCAGA GTGAGCCGGA GCTGAAGCTG GAAAGTGTGG TGATTGTCAG
1921 TCGTCATGGT GTGCGTGCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCAGAG
1981 CGCATGGCCA ACCTGGCCGG TAAAACTGGG TTGGCTGACA CCGCGCGGTG GTGAGCTAAT
2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCGAAAAA
2101 GGGCTGCCCC CAGTCTGGTC AGGTCGCGAT TATTGCTGAT GTCGACGAGC GTACCCGTAA
2161 AACAGGCGAA GCCTTCGCCG CCGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCCA
2221 GGCAGATACG TCCAGTCCCG ATCCGTTATT TAATCCTCTA AAAACTGGCG TTTGCCAACT
2281 GGATAACGCG AACGTGACTG ACGCGATCCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
2341 TACCGGGCAT CGGCAAACGG CGTTTCGCGA ACTGGAACGG GTGCTTAATT TTCCGCAATC
2401 AAACCTGTGC CTTAAACGTG AGAAACAGGA CGAAAGCTGT TCATTAACGC AGGCATTACC
2461 ATCGGAACCTC AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC
2521 AATGCTGACG GAGATATTTT TCCTGCAACA AGCACAGGGA ATGCCGGAGC CGGGGTGGGG
2581 AAGGATCACC GATTACACC AGTGGAACAC CTGCTAAGT TTGCATAACG CGCAATTTTA
2641 TTTGCTACAA CGCACGCCAG AGGTGCCCCG CAGCCGCGCC ACCCCGTTAT TAGATTTGAT
2701 CAAGACAGCG TTGACGCCCC ATCCACGCCA AAAACAGGCG TATGGTGTGA CATTACCCAC
2761 TTCAGTGCTG TTTATCGCCG GACACGATAC TAATCTGGCA AATCTCGCGC GCGCACTGGA
2821 GCTCAACTGG ACGCTTCCCG GTCAGCCGGA TAACACGCCG CCAGGTGGTG AACTGGTGTT
2881 TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGGATT CAGGTTTCGC TGGTCTTCCA
2941 GACTTTACAG CAGATGCGTG ATAAAACGCC GCTGTCATTA AATACGCCGC CCGGAGAGGT
3001 GAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGCAG GGCATGTGTT CGTTGGCAGG
3061 TTTTACGCAA ATCGTGAATG AAGCACGCAT ACCCGCTTGC AGTTTGTAAG GTATAAGGCA
3121 GTTATTGGTG CCCTTAAACG CCTGGTGCTA CGCCTGAATA AGTGATAATA AGCGGATGAA
3181 TGGCAGAAAT TCGCCGGATC TTTGTGAAGG AACCTTACTT CTGTGGTGTG ACATAATTGG
3241 ACAAACCTACC TACAGAGATT TAAAAAACCT CCCACACCTC CCCCTGAACC TGAAACATAA
3301 AATGAATGCA ATTGTTGTTG TTAACCTGTT TATTGCAGCT TATAATGGTT ACAAATAAAG
3361 CAATAGCATC ACAAATTTCA CAAATAAAGC ATTTTTTTCA CTGCATTCTA GTTGTGGTTT
3421 GTCCAACTC ATCAATGTAT CTTATCATGT CTGGATCGAT CCCCGGGTAC CGAGCTCGAA
3481 TTCGTAATCA TGGTCATAGC TGTTTCCTGT GTGAAATTGT TATCCGCTCA CAATTCCACA
3541 CAACATACGA GCCGGAAGCA TAAAGTGTA AGCCTGGGGT GCCTAATGAG TGAGCTAACT
3601 CACATTAATT GCGTTGCGCT CACTGCCCGC TTTCCAGTCG GGAAACCTGT CGTGCCAGCT
3661 GCATTAATGA ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTTCCGC
3721 TTCCTCGCTC ACTGACTCGC TGCCTCGGTG CGTTCGGCTG CCGCGAGCGG TATCAGCTCA
3781 CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG
3841 AGCAAAAGGC CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA
3901 TAGGCTCCGC CCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA
3961 CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC
4021 TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC
4081 GCTTTCTCAA TGCTCAGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTT GCTCCAAGCT
4141 GGGCTGTGTG CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATCG
4201 TCTTGAGTCC AACC CGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG
4261 GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA
4321 CGGCTACACT AGAAGGACAG TATTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG
4381 AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT
4441 TGTTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT
4501 TTCTACGGG TCTGACGCTC AGTGGAACGA AAATCACGT TAAGGGATTT TGGTCATGAG
4561 ATTATCAAAA AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT
4621 CTAAAGTATA TATGAGTAAA CTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC
4681 TATCTCAGCG ATCTGTCTAT TTCGTTTCAT CATAGTTGCC TGACTCCCCG TCGTGTAGAT
4741 AACTACGATA CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC
4801 ACGCTACCCG GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG
4861 AAGTGGTCCT GCAACTTTAT CCGCTTCCAT CCAGTCTATT AATTGTTGCC GGAAGCTAG
4921 AGTAAGTAGT TCGCCAGTTA ATAGTTTGGC CAACGTTGTT GCCATTGCTA CAGGCATCGT
4981 GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTACAGCTC GGTTCCTAAC GATCAAGGCG
5041 AGTTACATGA TCCCCATGT TGTGCAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT
5101 TGTGAGAGT AAGTTGGCCG CAGTGTATC ACTCATGTT ATGCGAGCAC TGCATAATTC

```

Figure 20 (continued):

```
5161 TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC
5221 ATTCTGAGAA TAGTGTATGC GGCGACCGAG TTGCTCTTGC CCGGCGTCAA TACGGGATAA
5281 TACCGCGCCA CATAGCAGAA CTTTAAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGGCG
5341 AAAACTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC
5401 CAACTGATCT TCAGCATCTT TTA CTTCAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG
5461 GCAAAATGCC GCAAAAAAGG GAATAAGGGC GACACGAAA TGTTGAATAC TCATACTCTT
5521 CCTTTTTCAA TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT
5581 TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGCC
5641 ACCTGACGTC TAAGAAACCA TTATTATCAT GACATTAACC TATAAAAATA GGCGTATCAC
5701 GAGGCCCTTT CGTCTCGCGC GTTTCGGTGA TGACGGTGAA AACCTCTGAC ACATGCAGCT
5761 CCCGGAGACG GTCACAGCTT GTCTGTAAGC GGATGCCGGG AGCAGACAAG CCCGTCAGGG
5821 CGCGTCAGCG GGTGTTGGCG GGTGTCGGGG CTGGCTTAAC TATGCGGCAT CAGAGCAGAT
5881 TGTACTGAGA GTGCACCATA TGCGGTGTGA AATACCGCAC AGATGCGTAA GGAGAAAATA
5941 CCGCATCAGG CGCCATTCGC CATTCAGGCT GCGCAACTGT TGGGAAGGGC GATCGGTGCG
6001 GGCCTCTTCG CTATTACGCC AGCTGGCGAA AGGGGGATGT GCTGCAAGGC GATTAAGTTG
6061 GGTAACGCCA GGGTTTTCCC AGTCACGACG TTGTAAAACG ACGGCCAGTG CCAAGC
```

//

Figure 21: Nucleic acid sequence of the known segment of the R15/appa transgene used for the generation of transgenic mice (SEQ ID NO:5).

LOCUS R15/appa 3470 bp DNA SYN 15-APR-2000
 DEFINITION R15/appa transgene with vector sequences removed.
 ACCESSION R15/appa
 REFERENCE 1 (bases 1 to 3470)
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 KEYWORDS salivary proline-rich protein, acid glucose-1-phosphatase; appA
 gene; periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

DEFINITION Rat salivary proline-rich protein (RP15) gene.
 ACCESSION M64793 M36414
 VERSION M64793.1 GI:206711
 SOURCE Rat (Sprague-Dawley) liver DNA.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1748)
 AUTHORS Lin, H.H. and Ann, D.K.
 TITLE Molecular characterization of rat multigene family
 encoding
 proline-rich proteins
 JOURNAL Genomics 10, 102-113 (1991)
 MEDLINE 91257817
 FEATURES Location/Qualifiers
 source 1..1748
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 /tissue_lib="cosmid genomic library"
 misc_feature 1802-1810
 /function=" consensus sequence for initiation in
 higher eukaryotes "

FEATURES Location/Qualifiers

DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
 gene,

ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
 VERSION M58708.1 GI:145283
 SOURCE Escherichia coli DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (bases 1811..3109)
 AUTHORS Dassa, J., Marck, C. and Boquet, P.L.
 TITLE The complete nucleotide sequence of the Escherichia coli gene appA
 reveals significant homology between pH 2.5 acid phosphatase
 and glucose-1-phosphatase

Figure 21 (continued):

JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE 90368616

```

FEATURES                      Location/Qualifiers
    Source                     1811..3109
                                /organism="Escherichia coli"
                                /db_xref="taxon:562"
    sig_peptide                 1811..1876
                                /gene="appA"
    CDS                         1811..3109
                                /gene="appA"
                                /standard_name="acid phosphatase/phytase"
                                /transl_table=11
                                /product="periplasmic phosphoanhydride phosphohydrolase"
                                /protein_id="AAA72086.1"
                                /db_xref="GI:145285"

/translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRA
TKATQLMQDVTTPDAWPTWPFVKLGWLTTPRGGELIAYLGHYQRQLVADGLLAKKGCPQS
GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLENPLKTGVCQLDNA
NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
YLLQRTPEVARSRATPLLDLIKALTTPHPPQKQAYGVTLPSTSVLFIAGHDTNLANLGG
ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIVSLVFQTLQOMRDKTPLSLNT
PPGEVKLTLGACEERNAQGMCSLAGFTQIVNEARIPACSL"
    mat_peptide                 1877..3106
                                /gene="appA"
                                /product="periplasmic phosphoanhydride phosphohydrolase"

    mutation                    replace(1817..1819,"gcg changed to gcc")
                                /gene="appA"
                                /standard_name="A3 mutant"
                                /note="created by site directed mutagenesis"
                                /phenotype="silent mutation"
    mutation                    replace(3092..3094,"ccg changed to ccc")
                                /gene="appA"
                                /standard_name="P428 mutant"
                                /note="created by site directed mutagenesis"
                                /phenotype="silent mutation"
    mutation                    replace(3095..3097,"gcg changed to gct")
                                /gene="appA"
                                /standard_name="A429 mutant"
                                /note="created by site directed mutagenesis"
                                /phenotype="silent mutation"

    polyA_signal                3262..3457
                                /note="SV40 signals"

BASE COUNT      1065 a      721 c      735 g      949 t
ORIGIN
1  GGATCCCCTT TGCTATGTAG TTTTAAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
61 GAGAGTCCTG TTTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
121 CTCTTTGTTT CTAGCATAAC CAAAAGATTT AGTGAATTGA AAACAATGTT CCCTTAGAGT
181 ATAGGTCTAA TAACCCCGAA AATATTACCA TGATACTGAG CATTTGTAAG TATCTCATAG
241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTTAA CATGATTTTC ATTAATCAGG

```

Figure 21 (continued):

```

301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGCACA
361 TATTTCACTA AACTAGGTTT ATCTATTTTG TTGCTTTCTC TAACATCTCT GCAATGAAGC
421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAT TTTCTACATA
481 TATCCTGGTT AGAGAGTGCT TAAAATAAGT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT
601 TGGGAAGAAA CCATTTGGTG AACAAATTTT CAAATAAAAA TAGACAAACA TAGTTAATTG
661 TAAAACATAT GTTTGACCAG CCCTTCTTTT CAATAGGCTT AATGTGAATA AAATGTTAAA
721 GATTCTCTTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT
841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT
901 TAAGATAAAG GTAACGTAT ACATTTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
961 TTCAGCTCTA TAATTCTTGC CTTAAACAAC TTAAATAGAA TGATTTAAAA TATGGAGCTG
1021 TCCATGGACC TTTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCCTT
1081 GATACTAACA CAGGTAAATC CCACACGTGT TTTGAGACTA CATTGCTGAG GATTTTATTG
1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTTGCTGTT ATTCGGTTAC TTCAAGAGAA
1201 TAGTGGCAAC TGATGCTATG TATCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATTT
1261 GTAAAGAAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
1321 GTGTTTAAGC TGTACTATTG ATCAAAGAAA TTTATTACCT TCAGTTTCAA TGGAAATAAT
1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAATCCTA TCCAAATGCA CAGTGATACA
1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTTA TCAATCAATT
1501 GTATGTATCA ATATATGGGC TATTTTCTTA CACATGATT TATTCAAATT TACTCTAATC
1561 ATTGTTGAAC CATTTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGGC
1621 AAAAGTCCCA GTGTGGAGTA AAGGATGCAA GATTTCTGCT TCTGTTAAGT ATAAAATAAT
1681 AGTATGAATT CAAAGGTGCC ATTCTTCTG TTCTAGTTAT AAAGGCAGTG CTTGCTTCTT
1741 CCAGCACAGA TCTGGATCTC GAGGAGCTTG CCGAGATTTT CAGGAGCTAA GGAAGCTAAA
1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTTTATCT CTTCTGATTC CGTTAACCCC
1861 GCAATCTGCA TTCGCTCAGA GTGAGCCGGA GCTGAAGCTG GAAAGTGTGG TGATTGTCAG
1921 TCGTCATGGT GTGCGTGCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCGAGA
1981 CGCATGGCCA ACCTGGCCCG TAAACTGGG TTGGCTGACA CCGCGCGGTG GTGAGCTAAT
2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCGAAAAA
2101 GGGCTGCCCG CAGTCTGGTC AGGTGCGGAT TATTGCTGAT GTCGACGAGC GTACCCGTAA
2161 AACAGGCGAA GCCTTCGCCG CCGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCCA
2221 GGCAGATACG TCCAGTCCCG ATCCGTTATT TAATCCTCTA AAAACTGGCG TTTGCCAATT
2281 GGATAACGCG AACGTGACTG ACGCGATCCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
2341 TACCGGGCAG CGGCAAACGCG CGTTTCGCGA ACTGGAACGG GTGCTTAATT TTCCGCAATC
2401 AAACCTGTGC CTTAAACGTG AGAAAACAGGA CGAAAGCTGT TCATTAACGC AGGCATTACC
2461 ATCGGAATC AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC
2521 AATGCTGACG GAGATATTTT TCCTGCAACA AGCACAGGGA ATGCCGGAGC CGGGGTGGGG
2581 AAGGATCACC GATTACACAC AGTGGAACAC CTTGCTAAGT TTGCATAACG CGCAATTTTA
2641 TTTGCTACAA CGCACGCCAG AGGTTGCCCG CAGCCGCGCC ACCCCGTTAT TAGATTGAT
2701 CAAGACAGCG TTGACGCCCC ATCCACCGCA AAAACAGGCG TATGGTGTGA CATTACCCAC
2761 TTCAGTGCTG TTTATCGCCG GACACGATAC TAATCTGGCA AATCTCGGCG GCGCACTGGA
2821 GCTCAACTGG ACGCTTCCCG GTCAGCCGGA TAACACGCCG CCAGGTGGTG AACTGGTGTT
2881 TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGGATT CAGGTTTCGC TGGTCTTCCA
2941 GACTTTACAG CAGATGCGTG ATAAAACGCC GCTGTCATTA AATACGCCGC CCGGAGAGGT
3001 GAAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGCAG GGCATGTGTT CGTTGGCAGG
3061 TTTTACGCAA ATCGTGAATG AAGCACGCAT ACCCGCTTGC AGTTTGTAAG GTATAAGGCA
3121 GTTATTGGTG CCCTTAAACG CCTGGTGCTA CGCCTGAATA AGTGATAATA AGCGGATGAA
3181 TGGCAGAAAT TCGCCGGATC TTTGTGAAGG AACCTTACTT CTGTGGTGTG ACATAATTGG
3241 ACAAACCTACC TACAGAGATT TAAAAACCT CCCACACCTC CCCCTGAACC TGAAACATAA
3301 AATGAATGCA ATTGTTGTTG TTAACCTGTT TATTGCAGCT TATAATGGTT ACAAATAAAG
3361 CAATAGCATC ACAAATTTCA CAAATAAAGC ATTTTTTTCA CTGCATTCTA GTTGTGGTTT
3421 GTCCAAACTC ATCAATGTAT CTTATCATGT CTGGATCGAT CCCCAGGTAC

```

//

Figure 22: Nucleic acid sequence of the SV40/APPA+intron plasmid (SEQ ID NO:6).

LOCUS SV40/APPA 5421 bp DNA CIRCULAR SYN 14-APR-2000
 DEFINITION Ligation of SV40 promoter/enhancer into CAT/APPA+intron
 ACCESSION SV40/APPA
 REFERENCE 1 (bases 1 to 5421)
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 KEYWORDS SV40 promoter/enhancer, acid glucose-1-phosphatase; appA gene;
 periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
 gene,

ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
 VERSION M58708.1 GI:145283
 SOURCE Escherichia coli DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (bases 40 1337)
 AUTHORS Dassa, J., Marck, C. and Boquet, P.L.
 TITLE The complete nucleotide sequence of the Escherichia coli gene appA
 reveals significant homology between pH 2.5 acid phosphatase
 and glucose-1-phosphatase
 JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
 MEDLINE 90368616

FEATURES Location/Qualifiers
 Source 40 1337
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 sig_peptide 40..105
 /gene="appA"
 CDS 40 1337
 /gene="appA"
 /standard_name="acid phosphatase/phytase"
 /transl_table=11
 /product="periplasmic phosphoanhydride phosphohydrolase"
 /protein_id="AAA72086.1"
 /db_xref="GI:145285"

/translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
 TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS
 GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
 NVTDAILSRAGGSIAFTGHRQTAFRELERVNLFPQSNLCLKREKQDESCSLTQALPS
 ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
 YLLQRTPEVARSRATPLLDLIKALTTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG
 ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIVQVSLVFQTLQOMRDKTPLSLNT
 PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"
 mat_peptide 106 1334
 /gene="appA"

Figure 22 (continued):

```

        /product="periplasmic phosphoanhydride phosphohydrolase"

mutation    replace(46.. 48,"gcg changed to gcc")
            /gene="appA"
            /standard_name="A3 mutant"
            /note="created by site directed mutagenesis"
            /phenotype="silent mutation"
mutation    replace(1320..1322," ccg changed to ccc")
            /gene="appA"
            /standard_name=" P428 mutant"
            /note="created by site directed mutagenesis"
            /phenotype=" silent mutation "
mutation    replace(1323..1325," gcg changed to gct")
            /gene="appA"
            /standard_name=" A429 mutant"
            /note="created by site directed mutagenesis"
            /phenotype=" silent mutation "

DEFINITION  Plasmid pBLCAT3  (bases 2200 to 4924)
ACCESSION   X64409
VERSION     X64409.1  GI:58163
SOURCE      synthetic construct.
ORGANISM    synthetic construct
            artificial sequence.
REFERENCE   1  (bases 2200 to 4924)
AUTHORS     Luckow,B.H.R.
TITLE       Direct Submission
JOURNAL     Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res
            Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG
REFERENCE   2  (bases 2200 to 4924)
AUTHORS     Luckow,B. and Schutz,G.
TITLE       CAT constructions with multiple unique restriction sites
for
regulatory  the functional analysis of eukaryotic promoters and
            elements
            JOURNAL   Nucleic Acids Res. 15 (13), 5490 (1987)
            MEDLINE   87260024
            COMMENT    Promoterless CAT vector for transient transfection
experiments with eukaryotic cells. Allows the analysis of foreign
            promoters and enhancers.
FEATURES
    source   Location/Qualifiers
            2200 to 4924
            /organism="synthetic construct"
            /db_xref="taxon:32630"

    SV40 t intron  1380..1993
                  /note="SV40 signals"
    polyA_signal  1990..2230
                  /note="SV40 signals"
    CDS           complement(3471..4317)
                  /codon_start=1
                  /transl_table=11
                  /gene="Amp"
                  /product="beta-lactamase"
                  /protein_id="CAA45753.1"
                  /db_xref="GI:58165"

```

Figure 22 (continued):

SV40 promoter/enhancer 5023..5402
/note="SV40 signals"

BASE COUNT	1413 a	1321 c	1331 g	1355 t	
ORIGIN					
1	CGAGATTTTC	AGGAGCTAAG	GAAGCTAAAA	GCCGCCACCA	TGAAAGCCAT CTTAATCCCA
61	TTTTTATCTC	TTCTGATTCC	GTTAACCCCG	CAATCTGCAT	TCGCTCAGAG TGAGCCGGAG
121	CTGAAGCTGG	AAAGTGTGGT	GATTGTCACT	CGTCATGGTG	TGCGTGCTCC AACCAAGGCC
181	ACGCAACTGA	TGCAGGATGT	CACCCCAGAC	GCATGGCCAA	CCTGGCCGGT AAAACTGGGT
241	TGGCTGACAC	CGCGNGGTGG	TGAGCTAATC	GCCTATCTCG	GACATTACCA ACGCCAGCGT
301	CTGGTAGCCG	ACGGATTGCT	GGCGAAAAAG	GGCTGCCCCG	AGTCTGGTCA GGTTCGGATT
361	ATTGCTGATG	TCGACGAGCG	TACCCGTAAA	ACAGGCGAAG	CCTTCGCCCG CGGGCTGGCA
421	CCTGACTGTG	CAATAACCGT	ACATAACCCG	GCAGATACGT	CCAGTCCCGA TCCGTTATTT
481	AATCCTCTAA	AAACTGGCGT	TTGCCAACTG	GATAACGCGA	ACGTGACTGA CGCGATCCTC
541	AGCAGGGCAG	GAGGGTCAAT	TGCTGACTTT	ACCGGGCATC	GGCAAACGGC GTTTTCGCGAA
601	CTGGAACGGG	TGCTTAATTT	TCCGCAATCA	AACTTGTGCC	TTAAACGTGA GAAACAGGAC
661	GAAAGCTGTT	CATTAACGCA	GGCATTACCA	TCGGAACTCA	AGGTGAGCGC CGACAATGTC
721	TCATTAACCG	GTGCGGTAAG	CCTCGCATCA	ATGCTGACGG	AGATATTTCT CCTGCAACAA
781	GCACAGGGAA	TGCCGGAGCC	GGGGTGGGGA	AGGATCACCG	ATTCACACCA GTGGAACACC
841	TTGCTAAGTT	TGCATAACGC	GCAATTTTAT	TTGCTACAAC	GCACGCCAGA GGTTCGCCCG
901	AGCCGCGCCA	CCCCGTTATT	AGATTTGATC	AAGACAGCGT	TGACGCCCCA CCACCGCAAA
961	AACAGGCGTA	TGGTGTGACA	TTACCCACTT	CAGTGCTGTT	TATCGCCCGA CACGATACTA
1021	ATCTGGCAAA	TCTCGGCGGC	GCACTGGAGC	TCAACTGGAC	GCTTCCCGGT CAGCCGGATA
1081	ACACGCCGCC	AGGTGGTGAA	CTGGTGTGTT	AACGCTGGCG	TCGGCTAAGC GATAACAGCC
1141	AGTGGATTCA	GGTTTCGCTG	GTCTTCCAGA	CTTTACAGCA	GATGCGTGAT AAAACGCCGC
1201	TGTCATTAAA	TACGCCGCC	GGAGAGGTGA	AACTGACCCT	GGCAGGATGT GAAGAGCGAA
1261	ATGCGCAGGG	CATGTGTTCC	TTGGCAGGTT	TTACGCAAAT	CGTGAATGAA GCACGCATAC
1321	CCGCTTGACG	TTTGTAAGGC	AGTTATTGGT	GCCCTTAAAC	GCCTGGTGCT ACGCCTGAAT
1381	AAGTGATAAT	AAGCGGATGA	ATGGCAGAAA	TTCGCCGGAT	CTTTGTGAAG GAACCTTACT
1441	TCTGTGGTGT	GACATAATTG	GACAAACTAC	CTACAGAGAT	TTAAAGCTCT AAGGTAAATA
1501	TAAAATTTTT	AAGTGTATAA	TGTGTTAAAC	TACTGATTCT	AATTGTTTGT GTATTTTAGA
1561	TTCCAACCTA	TGGAAGTGAT	GAATGGGAGC	AGTGGTGGAA	TGCCTTTAAT GAGGAAAACC
1621	TGTTTTTGCTC	AGAAGAAATG	CCATCTAGTG	ATGATGAGGC	TACTGCTGAC TCTCAACATT
1681	CTACTCCTCC	AAAAAAGAAG	AGAAAGGTAG	AAGACCCCAA	GGACTTTCCT TCAGAATTGC
1741	TAAGTTTTTT	GAGTCATGCT	GTGTTTAGTA	ATAGAACTCT	TGCTTGCTTT GCTATTTACA
1801	CCACAAAGGA	AAAAGCTGCA	CTGCTATACA	AGAAAATTAT	GGAAAAATAT TCTGTAACCT
1861	TTATAAGTAG	GCATAACAGT	TATAATCATA	ACATACTGTT	TTTTCTTACT CCACACAGGC
1921	ATAGAGTGTC	TGCTATTAAT	AACTATGCTC	AAAAATTGTG	TACCTTTAGC TTTTAAATTT
1981	GTAAAGGGGT	TAATAAGGAA	TATTTGATGT	ATAGTGCCCT	GACTAGAGAT CATAATCAGC
2041	CATACCACAT	TTGTAGAGGT	TTTACTTGCT	TTAAAAAACC	TCCCACACCT CCCCCTGAAC
2101	CTGAAACATA	AAATGAATGC	AATTGTTGTT	GTAACTTGT	TTATTGCAGC TTATAATGGT
2161	TACAAATAAA	GCAATAGCAT	CACAAATTTT	ACAAATAAAG	CATTTTTTTC ACTGCATTCT
2221	AGTTGTGGTT	TGTCCAAACT	CATCAATGTA	TCTTATCATG	TCTGGATCGA TCCCCGGGTA
2281	CCGAGCTCGA	ATTCGTAATC	ATGGTCATAG	CTGTTTCCTG	TGTGAAATTG TTATCCGCTC
2341	ACAATTTCCAC	ACAACATACG	AGCCGGAAGC	ATAAAGTGTA	AAGCCTGGGG TGCCTAATGA
2401	GTGAGCTAAC	TCACATTAAT	TGCGTTGCGC	TCAGTGCCCG	CTTTCCAGTC GGGAAACCTG
2461	TCGTGCCAGC	TGCATTAATG	AATCGGCCAA	CGCGCGGGGA	GAGGCGGTTT GCGTATTGGG
2521	CGCTCTTCCG	CTTCCTCGCT	CACTGACTCG	CTGCGCTCGG	TCGTTCCGGT GCGGCGAGCG
2581	GTATCAGCTC	ACTCAAAGGC	GGTAATACGG	TTATCCACAG	AATCAGGGGA TAACGCAGGA
2641	AAGAACATGT	GAGCAAAAGG	CCAGCAAAAG	GCCAGGAACC	GTAAAAAGGC CGCGTTGCTG
2701	GCGTTTTTCC	ATAGGCTCCG	CCCCCTGAC	GAGCATCACA	AAAATCGACG CTCAAGTCAG
2761	AGGTGGCGAA	ACCCGACAGG	ACTATAAAGA	TACCAAGCGT	TTCCCTCGT AACGTCCTC
2821	GTGCGCTCTC	CTGTTCCGAC	CCTGCCGCTT	ACCGGATACC	TGTCCGCTT TCTCCCTTCG
2881	GGAAAGCGTG	CGCTTTCTCA	ATGCTACGCG	TGTAGGTATC	TCAGTTCGGT GTAGGTCGTT
2941	CGCTCCCAAGC	TGGGCTGTGT	GCACGAACCC	CCCGTTTCAGC	CCGACCGCTG CGCCTTATCC
3001	GGTAACTATC	GTCTTGAGTC	CAACCCGGTA	AGACACGACT	TATCGCCACT GGCAGCAGCC
3061	ACTGGTAACA	GGATTAGCAG	AGCGAGGTAT	GTAGGCGGTC	CTACAGAGTT CTTGAAGTGG

Figure 22 (continued):

```

3121 TGGCCTAACT ACGGCTACAC TAGAAGGACA GTATTTGGTA TCTGCGCTCT GCTGAAGCCA
3181 GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC CGCTGGTAGC
3241 GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA AAAAAGGATC TCAAGAAGAT
3301 CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGAAACG AAAACTCACG TTAAGGGATT
3361 TTGGTCATGA GATTATCAAA AAGGATCTTC ACCTAGATCC TTTTAAATTA AAAATGAAGT
3421 TTTAAATCAA TCTAAAGTAT ATATGAGTAA ACTTGGTCTG ACAGTTACCA ATGCTTAATC
3481 AGTGAGGCAC CTATCTCAGC GATCTGTCTA TTTCGTTTCAT CCATAGTTGC CTGACTCCCC
3541 GTCGTGTAGA TAACTACGAT ACGGGAGGGC TTACCATCTG GCCCCAGTGC TGCAATGATA
3601 CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA TAAACCAGCC AGCCGGAAGG
3661 GCCGAGCGCA GAAGTGGTCC TGCAACTTTA TCCGCCTCCA TCCAGTCTAT TAATTGTTGC
3721 CGGGAAGCTA GAGTAAGTAG TTCGCCAGTT AATAGTTTGC GCAACGTTGT TGCCATTGCT
3781 ACAGGCATCG TGGTGTACG CTCGTCGTTT GGTATGGCTT CATTCAGCTC CGGTTCCCAA
3841 CGATCAAGGC GAGTTACATG ATCCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCTFCGGT
3901 CCTCCGATCG TTGTGAGAAG TAAGTTGGCC GCAGTGTTAT CACTCATGGT TATGGCAGCA
3961 CTGCATAATT CTCTTACTGT CATGCCATCC GTAAGATGCT TTTCTGTGAC TGGTGAGTAC
4021 TCAACCAAGT CATTCTGAGA ATAGTGTATG CGGCGACCGA GTTGCTCTTG CCGGCGTCA
4081 ATACGGGATA ATACCGCGCC ACATAGCAGA ACTTTAAAG TGCTCATCAT TGGAAAACGT
4141 TCTTCGGGGC GAAAACTCTC AAGGATCTTA CCGCTGTTGA GATCCAGTTC GATGTAACCC
4201 ACTCGTGCAC CCAACTGATC TTCAGCATCT TTTACTTTCA CCAGCGTTTC TGGGTGAGCA
4261 AAAACAGGAA GCCTAAATGC CGCAAAAAG GGAATAAGGG CGACACGGAA ATGTTGAATA
4321 CTCATACTCT TCCTTTTTTCA ATATTATTGA AGCATTTATC AGGGTTATTG TCTCATGAGC
4381 GGATACATAT TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTCCGCG CACATTTCCC
4441 CGAAAAGTGC CACCTGACGT CTAAGAAACC ATTATTATCA TGACATTAAC CTATAAAAAT
4501 AGGCGTATCA CGAGGCCCTT TCGTCTCGCG CGTTTCGGTG ATGACGGTGA AAACCTCTGA
4561 CACATGCAGC TCCCGGAGAC GGTACACAGT TGTCTGTAAG CGGATGCCGG GAGCAGACAA
4621 GCCCGTCAGG GCGCGTCAGC GGGTGTGCGG GGGTGTGCGG GCTGGCTTAA CTATGCGGCA
4681 TCAGAGCAGA TTGTACTGAG AGTGCACCAT ATGCGGTGTG AAATACCGCA CAGATGCGTA
4741 AGGAGAAAAT ACCGCATCAG GCGCCATTCT CCATTACAGC TGCGCAACTG TTGGGAAGGG
4801 CGATCGGTGC GGGCCTCTTC GCTATTACGC CAGCTGGCGA AAGGGGGATG TGCTGCAAGG
4861 CGATTAAGTT GGGTAACGCC AGGGTTTTCC CAGTCACGAC GTTGTAATAAC GACGGCCAGT
4921 GCCAAGCTTT ACACTTTATG CTTCCGGCTC GTATGTTGTG TGGAATTGTG AGCGGATAAC
4981 AATTTACAC AGGAAACAGC TATGACCATG ATTACGAATT CGGCGCAGCA CCATGGCCTG
5041 AAATAACCTC TGAAAGAGGA ACTTGGTTAG GTACCTTCTG AGGCGGAAAG AACCAGCTGT
5101 GGAATGTGTG TCAGTTAGGG TGTGGAAAGT CCCCAGGCTC CCCAGCAGGC AGAAGTATGC
5161 AAAGCATGCA TCTCAATTAG TCAGCAACCA GGTGTGGAAG GTCCCCAGGC TCCCCAGCAG
5221 GCAGAAGTAT GCAAAGCATG CATCTCAATT AGTCAGCAAC CATAGTCCCG CCCCTAACTC
5281 CGCCCATCCC GCCCCTAACT CCGCCAGTT CCGCCCATTC TCCGCCCCAT GGCTGACTAA
5341 TTTTTTTTAT TTATGCAGAG GCCGAGGCCG CCTCGGCCTC TGAGCTATTC CAGAAGTAGT
5401 GAGGAGGCTC GAGGAGCTTG G

```

//

Figure 23. The nucleic acid sequence of the Lama2/APPA transgene used for the generation of transgenic mice and transgenic pigs (SEQ ID NO: 7)

LOCUS transgene 17732 bp DNA SYN 14-APR-2000
 DEFINITION Lama-appA cut XhoI..20623 to NotI..17732
 ACCESSION transgene
 KEYWORDS parotid secretory protein; acid glucose-1-phosphatase; appA
 gene;
 periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 cloning vector
 REFERENCE 1 (bases 1 to 17732)
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

FEATURES

DEFINITION M. musculus Psp gene for parotid secretory protein.
 ACCESSION X68699
 VERSION X68699.1 GI:53809
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 3777 to 5332;)
 AUTHORS Svendsen, P., Laursen, J., Krogh-Pedersen, H. and Hjorth, J.P.
 TITLE Novel salivary gland specific binding elements located in
 the PSP proximal enhancer core
 JOURNAL Nucleic Acids Res. 26 (11), 2761-2770 (1998)
 MEDLINE 98256451
 REFERENCE 2 (bases 7147 to 12653; 13952 to 17731)
 AUTHORS Mikkelsen, T.R.
 TITLE Direct Submission
 JOURNAL Submitted (07-OCT-1992) T.R. Mikkelsen, Department of
 Molecular Biology, University of Aarhus, CF Mollers Alle
 130, 8000 Aarhus, DENMARK
 REFERENCE 3 (bases 7147 to 12653; 13952 to 17731)
 AUTHORS Laursen J, Hjorth JP
 TITLE A cassette for high-level expression in the mouse salivary
 glands.
 JOURNAL Gene 1997 Oct 1;198(1-2):367-72
 MEDLINE 9370303

FEATURES

Location/Qualifiers

source 1.to 12653; 13952 to 17731
 /organism="Mus musculus"
 /strain="C3H/As"
 /db_xref="taxon:10090"
 /chromosome="2"
 /map="Estimate: 69 cM from centromere"
 /clone="Lambda YP1, Lambda YP3, Lambda YP7"
 /clone_lib="Lambda-PHAGE (Lambda L47.1)"
 /germline
 /note="Allele: b"

misc_feature 3777-5332
 /gene="PSP"
 /function="salivary gland specific positive acting
 regulatory region"
 enhancer 7147..8724

Figure 23 (continued):

```

                /evidence=experimental
exon           11778..11824
                /gene="Psp"
                /note="exon a"
                /number=1
                /evidence=experimental
exon           12626.. 14190
                /gene="Psp"
                /note="exon b fused with exons h and i"
misc_feature   12644-12652
                /function=" consensus sequence for initiation in higher
                eukaryotes ."
misc_feature   13952-13965
                /function=" M13mp18 polylinker"

```

DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene,

```

ACCESSION  M58708 L03370 L03371 L03372 L03373 L03374 L03375
VERSION    M58708.1  GI:145283
SOURCE     Escherichia coli DNA.
ORGANISM   Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision;
            Enterobacteriaceae;
            Escherichia.

```

```

REFERENCE   1  (bases 12653..13951)
AUTHORS    Dassa,J., Marck,C. and Boquet,P.L.
TITLE      The complete nucleotide sequence of the Escherichia coli
            gene appA reveals significant homology between pH 2.5
            acid phosphatase and glucose-1-phosphatase
JOURNAL    J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE    90368616

```

```

FEATURES             Location/Qualifiers
     Source                12653..13951
                           /organism="Escherichia coli"
                           /db_xref="taxon:562"
     sig_peptide           12653..12718
     /gene="appA"
     CDS                   12653             13951
                           /gene="appA"
                           /standard_name="acid phosphatase/phytase"
                           /transl_table=11
                           /product="periplasmic phosphoanhydride
                           phosphohydrolase"
                           /protein_id="AAA72086.1"
                           /db_xref="GI:145285"

```

```

/translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
TKATQLMQDVTTPDAWPTWPVKLGWLTTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS
GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS

```

Figure 23 (continued):

ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF

YLLQRTPEVARSRATPLLDLIKALTTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGG

ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIVSLVFQTLQQMRDKTPLSLNT

PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"

mat_peptide 12719 13948
/gene="appA"
/product="periplasmic phosphoanhydride
phosphohydrolase"

mutation replace(12659.. 12661,"gcg changed to gcc")
/gene="appA"
/standard_name="A3 mutant"
/note="created by site directed mutagenesis"
/citation={3}

mutation replace(13934..13936," ccg changed to ccc")
/gene="appA"
/standard_name=" P428 mutant"
/note="created by site directed mutagenesis"
/citation={3}

mutation replace(13937..13939," gcg changed to gct")
/gene="appA"
/standard_name=" A429 mutant"
/note="created by site directed mutagenesis"
/citation={3}

BASE COUNT 4719 a 4125 c 4168 g 4719 t
ORIGIN

```

1 TCGAGAGTAT CTTTGTGAGC TGTGCCTCCA ACAAAGGGGT ACTGTTGCCC ACATAGAAAG
61 ATCTAACTA ATTAATTAAT CCCTCACCCG CAAATCTTTC AGTCACTAAG TTAGCACGAT
121 TGTGAACAA GTTCTCCAAA GGAGAGATAC AGATGAGTGC GTATAGGGTG GACCTGGCTG
181 CTGAGGAGAC ACCTGCATCT GACTAAGAAG AGCCACGGTG TTAGTTGAAT GGTGTGGAGT
241 AGGGTGGTTC TGTGGGACAG TAGAAAATCG AGAGGCATGT GCCGTTTAGT GAACTGATGG
301 AAGCTACCCC AAACGACAGA GATTGTGAGT CAGGCCAATC CGTTTCGAGT TTGATGGGCA
361 GCCGGACAGT GAGACAGACA CACCTACTCA GTTGGAGGAA GGATGAGAAC AATGGCCAGC
421 AGGGATTGAG AGACCCTGAC AGGCGCAAGG CCCTAACACA CACACCTACC ACCTCACTTG
481 ACAAAGCTGC CAAAGACCAA AGACTTGTTT TCCATTAGAA ATGACAGCTG GCTTGACCCG
541 ACAGCATAAT AAGCAGAGTG TACTCTGATT GGAGAACTTT AATGTGTTTC ATTCAGTATT
601 ATAAAAGGAC AGTATTACAG ATTTTGTTGT ACACTGCTGT TACATGTGGG GCAGTGTGTC
661 TTTAAGTAGG GTAAAGTACT CTTTAAAAAT GGGTCCTAGA TATTTTTTCC TTAACTCAA
721 GTCTCTTACT GTTTAAATGA TTTTATTTT GTTTAATATG GAGGAAAAAG AAGCGTAAAT
781 GGACAATATA TATTTAGAGA AAGATGGTTA GCTGTCAGAA AAATATGCAA ATCAAAATCA
841 CACCAAGACT GCAGCACACC CCTGTGAGAT GGCTGTGATC AAGAAAATAA ATGACAATGA
901 GTGGTGGTGA AGATGTACTA AAGGGAAACA CACACACACA CACACACACA CACACACACA
961 CACACTGGAG CAACCACTGT GGAAATCAGT ATGAATGGTC CTCAAAAACC TGAAGATAGA
1021 GCGGGGCGTG GTGGCATAACA CTTTATTTC CAGCACTGGG GAGGCAGAGG CAGGTGGATC
1081 TCTGAGTTCC AGGCCAGCCT GGTCTATAGC ACAGGTTCTA GGACAGCCAG GGCTACACAG
1141 AAAAAACCTG CCTTGATTAA ACCAAACCAA ACCAAACCAA ACCAAACCAA ACCAAACCAA
1201 ACCAAACCAA ACCAAACCAA ACCAAACCAA AACACTGAAG ATAGAACTTC AGTATTCCAT
1261 TCCTAGATAT ATACCCAATG GAGACTAAGT CAGCAAGACA CCTGCACAGC CATGTTCACT
1321 ACTACACTGT TCACCACAGC CAGGCTGTGG AACCAGCCTG AGTSTCCATG ATAAATGAAT

```

Figure 23 (continued):

1381	GGATAGGTAA	CTTCAAGGT	AAATGGACTC	TGCTGTGTAC	ATGCCTCACA	TTCTGTTTAT
1441	TCATTTTTCT	TTATGAGGTG	TCCATTCAGG	AGTCACATGG	TAGTTCTATT	TTCAGTCTTC
1501	TGAAGATACT	ACACTGGTCC	CCACAGTTTA	CACTTTTATC	AGCAGTGAAT	AAGGGTTCCT
1561	CTATCCTTAC	CATCATTTGT	TGTAATTTTT	CTTGATGACC	CTCTTTCTGA	CAGGGATAGG
1621	ATGTAATATC	AGTGTGAGGA	AGTACAACCT	GTTTTCTAAG	TATTTATTGG	CCCCTTGCAT
1681	TTCTTCTTTT	GAAAACTGTC	GGTTCCTGAC	ATCTGCTCAG	GTATTCATTG	GATGTTGTTT
1741	CTTTGGTGTT	TGAGTTCCTA	TGAATTCCTAG	ATGTTAAATC	CCTGCCTGTG	GTTCTCTCCC
1801	ATTCTGTAGG	CTGCCTCCTC	ACCCTGGCAA	TTGTTGTCCT	TGTTTTGCAG	AAACTTTTGA
1861	CTTCATGGAA	TCTCATTTGT	CAGTTTTCCC	TCCTCTGCTA	TAGCCTGAGC	TAATGCACTG
1921	GTTTTTACAG	AGCCCTGGTC	TATGCCTTTA	TCCTCCTCTG	GCAGCTTCGG	AGTTTCATTT
1981	CTTACATTTA	GATCTTTGAT	CCACTTTGAA	CAAGTTTTGG	AGCAGGGTGA	GAGATACGAA
2041	TCTAGTTCCA	TTCTTCCATA	TGTGATCCTA	GTTTACATAG	CATCGTTGGT	TGAAGAGGTT
2101	TTATTTTATT	TTTAAATAAT	GTGTCATAAA	AAACGAGGTG	GTTGTAGCAG	TGTGGATTGG
2161	TTTCTTTGTC	CTTTGATCTA	CAGGTCTTGT	TTTGTGTCAG	TCTCATGATG	TTTTATTGCT
2221	ATGGCTCTGT	CATACAGTCT	GAGGTCAGGT	ATTGTGATAT	ACCTTCAGTA	TTGCTCCCTC
2281	AGACTCAGGT	TTGCTTTGGC	CAGGAGTCAT	CTTACTCAGT	GCTCTTAGAG	CTCCCCCAGC
2341	ATGTAGCTGC	TACTATTCTT	AGTTGATAAA	TCAGGAAACT	GGGGCTCAGA	GAGATTAACT
2401	GTCTTGAAGT	ACTTCTGGGG	AGGTGAAACG	TGGAGACACT	AAACTGTGTT	TACCCTGTAC
2461	TGCTCCAGTA	GCTGTCGGGT	GCTGGGCTAC	AGCAAAGCAC	CTATACTATA	TATTACTCAG
2521	GAGGTGGAAA	AACTCAGCCT	CCCTTGGGGT	TCCCAAGCTC	CCAGGTGTCC	AGTCACTGCT
2581	GGAAACCTCA	TGGAGTCTGA	AAGGAAGGGT	TGAGGGTACA	TGGGGCAGCG	ATGAGGAGCC
2641	TGGGGCTGGG	ATCTCCCAAA	CACCTGGATA	TCCAGATGCC	ACTGGGTCAG	GGGGAGTTGG
2701	GAACAGAGTT	GGGATGTCCA	TGGACCTGTG	ACAAGGCCAG	GGCCAGGGGG	AGGATAACTC
2761	TGGCTTTACT	AATTTGCGAA	AGTCCTTAGC	TTAGCAGCAG	TTGTCTGGGA	GCACAGAGGG
2821	GCCTTCTGTA	AGAGGCTCAG	GCAGTGCCCG	TCTGTAGGCG	AAGGTCTTCT	CCATGTTCCC
2881	CATGGTGGTT	CTTGATGAAA	GAGACAGTCC	TTGGCTCCAA	ACTGGTTTAT	TGATTGTTCA
2941	TTGTGGAAAA	TGGGTGCACA	CCACCTTCTC	AGGGTGGACC	AGAGATCAAA	TACCTTTTGC
3001	AGGGAGGAAT	ATCTGGGAAG	GGACGCTTAC	TGGCTAAACC	CTCAGGGCCT	CTAGATACAT
3061	CATTAGCATG	GAGAACTCTG	TTCTGGGCTA	CATGACCACA	GGCCACATTT	CCACAAGCCA
3121	CATGTGGGAA	GTGTGGCACA	TGTTCTAGGC	CAGGAATCTG	GTAGGGAGCG	TGGAGCCACC
3181	TACCATCCCA	GGTGGGTGCC	TGGGTGCCAG	GGACCTTGAA	CCCGCTCAAC	CTTACCAAGT
3241	TTCTTGGCAG	GGTCCACTGT	CCTACACAGA	AGCTGGAGGA	GGTGTGAGGG	TTGTGTCTTT
3301	GTGGAATGTC	CCATGCTGCT	TGGGGCTCAG	TTTCTCCACC	TGTACCTCAT	TGTTTGGGT
3361	ATAAAAAGTG	GGGATACTTT	ATTATTCTCT	GACTCGGTCC	TGAGGAAAAA	GACTCGTGCC
3421	AGTCCAGGAA	CCACACCCTG	AGGTTCTCTG	ACTGAAGGGA	CTCCCTAAGT	CTCTGGAGTC
3481	TCTCCCTTTC	ACAGAGCTGC	CAAAGTCTAG	GTTCTTTTGA	GGATAACAGA	GCCATGCTTG
3541	GTAAGCAGAC	AACAGCATTT	GTTTACTCAA	CCTTCTTTTG	TCAGCTCCCT	CTTCATAAAC
3601	AAGTTGAGAC	ACCATGCTGG	CTTGAGGAAG	ACTTCTAAAG	CCAGACAACCT	GTGCAAGGAA
3661	GAAGAAGAAG	GGGCAAGTGG	AGTTAGCCTG	GATGTAGCCC	TCAAAGTCTC	CAGAGACCAG
3721	CCATGAAGGC	TCAAGTGGAG	GGCAAGACCT	GCAGCAGCCA	AGCATCTGGC	AGGAGAGGAT
3781	CCTGGGAACC	CCTCTACCAT	GACACACATT	CTTCCTGCAG	GTCACACTTA	ATAGGCCATT
3841	TCTTATTTGG	ATCTATCATG	GTGTTCTGTG	CGAGATTAAT	GAGGTGTTAT	GCTGCGAACA
3901	GAAAGTTATA	TAAAAACAAG	TCCCCCCCCC	TTGTCACTGC	TGCTAAGAAT	GTAGCAGAAA
3961	TTGTCTCAAG	TGTCTCTCTA	ATCAGAAACA	ATAAAGGTCT	CCTTGGATTTC	AAGCCCTCCA
4021	GTTTCCTCCT	TCCTTGCTGA	GCCTTGGACA	CCCATACAAA	CCTCCTGGAT	GCTACAGCTC
4081	TGGGCAGAGA	CTCCAAGGTG	GGGAGAGACT	GATGGTACAA	AAGCAAAATA	CTTGTTTGGG
4141	GGTACACCCA	CTCCTCTGCC	TGTGTGGTTC	CTGCAGTCAG	TCCTGCAGAC	AGGCCCTCAG
4201	TGGGTCTTCC	ATGGGCAACA	CGCAGAGGGA	GGCAATGGAT	GGGAATACCC	ACACCCTGGT
4261	TAGTTTACCC	CGGCCATGCT	CTCTGCTCTT	CATCCCTCCT	CTGCCCTCTG	CCACGGCTTT
4321	CTCTGCAGGA	ATCATATCTT	CATATTGGCC	CACAGGTGTT	CTCCTCACCC	TAGCTATGAT
4381	GTTTACTTTA	GAGTGACCTT	AGCAGGGCTG	GTGGGAATGA	GTTCTAGAAG	GCTCACGGAG
4441	ATGCTAGGGA	AGAAACGTCT	TCTAACTACT	GAGGTTACTA	AGTTCTTGTT	GGTTGTCTCT
4501	GCCTTTCCCT	TGTTAAAAGTC	ACCTTGAAGT	TAGTGCAGAA	GAAATCAGAG	CCAGTCACA
4561	GAGTAAATAT	GGTCCTGAAG	ATTTCCCTTG	AGTGCCCAAG	ATCCATGACA	TTTCAAGAGC
4621	CCTCTTTGTA	CCTTAAGTCA	TTTGGGGTTG	TATCTTCTGC	TTGATGTATG	TGTGTGTGTT
4681	TATCAAAGAG	TGAGATGGTT	ACATAAGAGG	TGCTCTAAAG	GACAGAGAGG	ATTTGCAATT
4741	GTGGCATGTG	ACATCCTCAG	GCCTTGCTCT	GGTGCCAGGA	GGAAGTGTATG	CAGAAAAGAG
4801	TAAGAGGTCA	TTTCCTGGAG	GCTGTCACTA	TAGAGGAGAT	CTTAGAGTGC	ATTCCTCCT

Figure 23 (continued):

4861	CCAGGCCCTG	CCTGAGGATA	GACATGTGCT	GAATGCAACT	GAAACAGAGG	CTTGGGATGG
4921	AGAGTTAGGT	TCACAGAAGG	GAGGGTGGGA	GATGGATGCT	TGCTGGGTTC	TGGGTCTCAT
4981	CACCAGCTCC	TGACCACCCG	GTCAGCCCAT	GTGCTTATTC	CATAGCTTTC	TTTTGCTATG
5041	TTTACTCAGT	GTGGTGTGTT	TTGGGACCCA	GCAGAAGCCA	GTCCCAGGCT	GACAGCTGTG
5101	GATACACAGG	GCAGCATGAG	GGTCCTCAGC	CTGAAGCAGT	CAGGCTGGCA	GAAGAGAAAG
5161	ACCAGCACAC	ATTCTTTCAA	CCAACATATG	CTTGAAAAAC	AAACATATTA	TATCACATAT
5221	ATTGCATTTA	TGAGACAGCT	AAAATGTACT	CGGGTAGCAT	GACTCCAGGT	GGGGATATCT
5281	GCAAGTGCCA	TGAGTGGCAG	AGGGACAGCC	AATGTGAGGC	AAGAAGGAAT	TCTGGCTCAA
5341	CACAGCTTAG	CTCCCTGGTG	TTGGTTCAAA	CTTTGAGAGT	TTGACCACAA	GCACTTTATT
5401	TTTGACATAT	TTAAACAGAG	CACAACCTTG	GGAAAAAGTT	TTCTTATGAA	AATTATCACA
5461	ATAAAGCTTA	AGGCATGACT	ACATTAAAAT	GCCTTTGCAA	AGTATATGTG	CCCTCTTCCA
5521	CAAGAATGGT	TCTATTGACT	GAGAAATAAT	GTTCAGGATA	AAGATCCAGG	AAGAAAAGAT
5581	CAGGGATAAG	TAAAATACTA	AACCTTTTTG	CAAAGTACAT	AGACCCTCTT	TCATAACAAT
5641	GGGTTCTATT	GACTGACAAG	CAGTGTCTAG	GAGTTGGGAA	AGAGTCTAGC	ATAAGCACGA
5701	TAGCCTGGAG	ACTCTAGTGA	GGTCTAGTCT	TACAGACAGC	AAAAATCACC	AGGTTACAAA
5761	CTACATTCAT	TTCCAGTTTT	CTGATCAGGC	ACAGGTATGA	ATCCCTTCTG	TTGAAGAGAA
5821	AAGTCCATGT	GTTTAAAATA	TCTGGTTTCT	CCAGTGCTAT	TAGCGAGAAG	ACTTGAGCCC
5881	TATACAACCT	CCACCTGGAG	TGACATCCTG	TCTTCATGGT	ATATTACATA	CCTAGACACG
5941	CTCATCTCAC	AGACTTAGGA	CTTTGTCTTC	TGATCTCCAT	TTCTGATCCC	ACTTCCACCT
6001	TTGCCTTGAT	AGTGTCATTT	TCTTCACTGC	CTTGGTGACA	ACCATGTTAT	CCTCTGTGTA
6061	TTTGAGTGTT	ACCATTTTCA	GATTTTACCT	GTATGCAAGA	TCACACAGTC	TTTGCTTTTC
6121	TGTCTGGATG	CATGCTAATC	TCTACACAAC	AACCCTTCCC	CGTCACTCAG	ATCTTCTCTC
6181	ATTAACACAT	ACATGGTGCT	GAAGAGGCTA	GGGAGCTTCC	CTTCAGTGGG	GAGCTAGCTG
6241	GCTATTGGGC	CTTTTTGACT	GTCCAGGAAG	GCCCCAATT	GCTGAGACAA	GAACCTAGAT
6301	TCTTCATTAT	TGACTCTAAC	TCATGTATCA	AGCAGAAGCT	AATGAATAGT	TATCAACAGG
6361	ATCAGAGGTT	CCAGTGTAAG	ACACTTTGAC	ATGAAAGAAC	GGAGGAAGGA	CAGATGGATG
6421	CATAAAAGCA	GGACCACTGC	CCCAGGAAGG	TCCTGGAAAC	TGATGCAGGG	CAAAGGACAG
6481	GTTATAAAACC	AAATCTTAGG	GAGTCAGGAA	GAGCACAGAG	GAGCTCAACC	AACTGACCAC
6541	TGCTTAGGGG	CTACCAACCC	AATCCTCCCT	GTGGGAACAG	CTAAGCTATC	AGCCAAGGGT
6601	AATAAACAGG	CAGGACCTGT	GGATGACATG	GAGAGCATAG	GGACCCTGGG	TCCAGCCTTT
6661	AGCACCTGCA	CTCTCAGGAT	ACTCCACCAT	TGTGTCTTAG	AGAGCCTAGG	GATACTGGGT
6721	CCAGCCTTTG	GTACCTTCAC	TCTCAGGGTA	CCCCATCACT	GTGTCTTGGG	GAGCCTAGGC
6781	ACCCTGGGTC	CAGCCTTCAG	TACCTGCGCT	CTCAGGACAC	CCCACCATTG	TCTCTTGCCC
6841	CGTCTCTTCT	TCCTCTTCTT	CCCTTTTCATT	GTCTCTTCTC	TGTTTCTTTC	TTGACTCTCC
6901	TTTCCCTTCA	CACCCTCACT	CTAGTTCTCC	CCTTCCCTCT	CTGCATCACC	CTATTCTCTC
6961	TGTGTCTCCCT	CCACTTTTCT	TTATCTCTCA	TGCTTCTCTC	CTCCCTCAAA	TACTTGTAC
7021	CCACTATACT	TCAGGGGCCA	GCTCTAGTGA	CAAAGCTGTT	AATAGCAAGA	CTCTCAGATC
7081	TCCAACGGCT	CAGAGGAGCC	AGACCCACCA	AGAAGCTCTT	CCAGGTCCAA	TTTCAGGTTT
7141	CTTCGAAAGC	TTTCAGCAAA	TGCTCAGGGA	ACATGCCACT	AACAAGAAGA	TGCAAATTCC
7201	AGTTGAGAGT	GGGAAAGGCC	CTTGCGTAGG	TCCCATCTTC	CAGGCCAAGG	TCAGAGGGGC
7261	TCTGTGTAAT	CCGGATTGAC	AGGGCTCAGA	ACAATGTTTT	GTTTTTAAGG	TTTATTTATT
7321	TTAGGTGTTA	GTGTCTTTGC	TTGCATGACC	TTATGTGCAT	CATGTGTGTG	CAGGTTCCCTG
7381	ATGACAGTAG	AGGAGGGCTT	TGAATCCCTG	GGGATAGGAA	GTTACAGGAA	ATTATAAGCT
7441	GCTTTGTGGG	TCTTCTAGCT	TTCCCAACAG	AAGTGAATGC	TCTTACCAC	TGAGCCATCT
7501	CTCTAGGCC	AAGAGACATT	GCTTTATGGA	TATAATTGTG	TGTGTGTGTC	AACATTGAGG
7561	AAAGGGAAAT	AAAAAAAAAA	CTTCAGCCGC	TAAGGTTGTA	CAGTTTCACT	AATTGCTACT
7621	TTTAGTTGTG	ATAAAATGGC	AGGTGCTTCA	ACATTTATAT	ATACAAAAAC	TTCCCTGCTG
7681	GTGGTTCAAC	TGTGAGAACT	GGGGTAAGTG	GGTGAGTTCT	CTTTTCTGT	CTCTGTCTCT
7741	GTCTCTCTCC	TTCCATTCTT	TCTTAAAGGA	AATAAACATT	GCAGCTGGGT	TATAGCTCAT
7801	CAATATGGAA	GTTACAGAAG	TGAAAAAAGG	CATTGCCTTG	GTGGGTGGTG	TTACCAGCTG
7861	ATTTTTGGTT	GTCCTGCAAG	GAGGTCTGGG	GACTGGCTGC	TCTGTCTCTG	TCTGTATGAG
7921	TGAGGGAAGT	CTGGGGAGCA	GATTCCCTAA	CCCTCAGCCT	GGCCTGGTTC	CTGAGTGAAC
7981	CCAGCCTCTC	TGGTCCTAGT	AGCTTTTTTC	AAACAGGAAT	CTGAGTGGTG	ACAGGGAACA
8041	AGTACCAGCC	CATTGCCTAA	GTGCCAGGGT	TAGTGAGGGC	AGGAAGCTGC	CATAGCTGGG
8101	ATTAGTAGTT	GTATTGGATG	TAGGAAGTCC	TATCCTGGGA	CAGCTAATCC	TTAATGCTTC
8161	ACTGGAGATT	TTCAATGAGA	AATTTATCCC	ACGGCCCATC	TGGCCCCATC	CTTTTGTCTC
8221	CAACAGCCAA	GTATTTTCCA	TTAGAGGAGA	CTTCCTGTAC	ACTTGATGGA	TGCTCATTC
8281	AAGGTGACTT	GGGGCAGTCA	GTACAGACTT	GGGATGACCT	CTGACAGGCT	AACCTCTCCC

Figure 23 (continued):

8341	CAACAAGGGC	CCTCTATGTT	TGCTATGTAA	TGTAATGTCA	GACATTGTCA	GGAGTGTCCC
8401	CAGCACAGCC	TGCCCAGTGT	GAGGGCTCTC	ATAGGTTTCC	CACTGTCTTA	TCTACACAGG
8461	GATAACGAGG	AGGTAAGCTG	CAGTTCCAG	TCTCACTTCA	CAGAGGAAGA	GATAACCCCA
8521	TCCCAGGTCA	TGTAGCCAGC	AGTGGAAAGA	ATGAGGATTT	GAACTCAGGT	CTTCCAAGTC
8581	CCATTGATAG	CATCTCCTCA	CAAGTCCCTT	GCCACCCTCA	CGATGCCTTA	GACACTTGCC
8641	TGCCCTTTAT	ACTAAGGAGA	TGCAGGTACA	AGGGGTTTAC	CCATGTAGCA	GCTGAGGCAG
8701	CTGGGGATAG	ATACCAGCAG	CAGGCCTGAT	GTCACCACTC	TAACCTCCAGC	ATCCCCAGTC
8761	TGTGTTCCCTG	GAGTGTGAAA	ATCCCTACTT	AACAAGATTG	TGCAACAGTC	CTTGCTCTG
8821	TGACCCATAG	CTGGAAACAG	GATTCTCATT	GATTTGTGGA	ACATGGTGGC	AGCCAGCCAA
8881	AAAGAGGGTC	TGCATACAGA	AGACACGTGT	GGCAAGGCCA	CAGCAGACTC	TGACTACCTT
8941	AGCTTACAGA	ATTACAAGGT	CATAATGTCC	TCTGCTTTGG	TCACCTCATG	TTAAGGACAG
9001	GCCCTAATGA	AGATGGGGCA	GAAGACTGAA	GGAATGGCCA	ACCAATAACT	GGCCCAACTT
9061	GAGACCCATC	CTACAGGCAA	GCATCAATTC	CTGACACTAC	TAATGATACT	CTGTTATGCT
9121	TGCAGACAGA	AGCCTAGCAT	AACTATCCTC	CGAGAGGTCC	ACCCAGCAAC	TGACTGAAAC
9181	AGAAAAAGAT	ATCCACAGGC	AAACAGTGGA	TGGAGGTCAG	GGACTATTAT	GGGAGAGCTG
9241	TGGGAAGGAT	TAAAAACCTT	GAAGGGGATA	GGAACCCAC	AGGAAGACCA	ACAGAGTCAA
9301	CTAAGAGACC	TGTGGGAGCT	CTCAGAGACT	GAGCCACCAA	CCAAAGAGCA	TACACAGGCC
9361	GGTCCGAGGC	ACCTGGCAGC	TGTGAAGCAG	ACATGCAGCT	CAGTCTCCAT	GTAGGTCTCTC
9421	CAATAAGCGG	TAGCCTGACT	GCAGTATCCA	ATCCCCAACA	GGGCTGCATA	GTCTGGCCTC
9481	AGTGGGGGAG	GATGCCCTTA	ATCCTGCAGA	GACTTGATGA	GTGGAGAGCT	ATCCAGGGGG
9541	AACCCACCCT	CTCTGAGAAG	GGAATGGGGA	TGGGGGAGGG	ACTCTGTGAA	GAGGGGACAA
9601	GGACAAACAA	GAACCTCAAA	TAGGTGAGGC	CCTAAAGGCT	TGCTAAGTAG	CAGTGGCCCA
9661	GCTCTGTCTT	GTTCTCTCAGC	CCAAGGCTCA	GCTCCACCTT	GTTTCTGTGT	TTTTCTGGCT
9721	TTTCATGGGC	CTAGGACTTG	GTGACCAGTT	CAACAATGG	GGCCTGTGGA	AGACACAATA
9781	TACAAGACTA	GGGACATTCC	TGTTCTGCTG	ACTATCCATA	GCCTGATGTA	GGTGGAAGGA
9841	CCCAATCACT	GGATTCTTAC	CCTTGACCAA	CCTTGACAGC	TGAGGGCCTC	TCAGAAACCT
9901	ATTTCTTCCA	CTGAAAAATG	AGACTCTCAA	ATGAACGTCG	TGACAATCAT	CAGGCTTATT
9961	AAAGAGGTGT	ATCTAACCTG	AATGGCAAGC	AGACAGCAGG	CAAATGTCTG	TATCAACCTC
10021	TAGGAAGGAC	AAGAACTGCT	CACTGCTGCC	CCCCAGGAGG	CCATTTGCTG	AAACAGCTGC
10081	TCTCCTGCTG	GTGCACAGGC	CCTGCCTTCT	CATTGCAGCC	ACAGCCCCTT	CCTGTCTGAA
10141	CCTCCTGTCA	GGTCACTGGG	AAACAGATCA	AGATGGAACA	GGACAGCTCC	TGATGGTAAA
10201	TAAAAAACAG	TGGTCATGGC	TATTCATAGG	GGTTTATGCT	TCTTCAGTCC	ACACTGTGAA
10261	GAGCTGTGGG	CATGAACCAC	AGTGTTCGAG	GTAGAGTTGG	GGTTCTGAAA	TTACAGTGG
10321	GGTGAGCTCA	GTAAATGTGA	GCTGGAGGTC	ACTCGTGAGA	CACACAGTCC	TGCTGCTTCT
10381	GTTCCCAATA	TCCTGAGGAG	ACGACACATC	TACTTTGTTC	AGAGGCCACA	GTCTAGTTGA
10441	CCTGAGAGTT	ACCAGTTTCT	TATTTGTGTG	TGTGTGTGTG	TGTGTGTGTG	TGTGTGTGTG
10501	TGTTGTTTCGT	GTGTGAGTGC	AGGTGCACAT	ATGATAGCGT	ACACGTTGAG	GTGAGAGGAT
10561	AACTATCAGG	CGTTGTCCCC	TCCTACTTTT	CCTCGGACTC	TGGAGAACAA	ACATGGGTCC
10621	TTATTCCAGG	GGAGCAAGTC	GCTGTTGGCT	GACACATCTT	GCTCACATAC	ATTTTACCTA
10681	GACAATGGAG	CCTCCATCAG	AGTATTACTT	TAGCTCCTCA	CCGATGGCAA	TGCACCACCT
10741	CTCTACCCAC	ATAGGAGTTG	GGTCTCCACA	CACCCCCACA	CCCCCTTCAC	CAAAACGTTT
10801	TCAGTTACTT	TATCTGGTAA	AGTTCATCAG	AGAATGAAGC	CAGTATTAAG	AACATGGAAT
10861	CATTTGGGAA	CCTGGATCTA	GCAATACCCC	ACCCTAGATG	GAGTTGCTGA	GTTTTACCTT
10921	CAGATTATAA	TTCCCCCCTA	GCTTCTATGG	TTTATTCTGA	AACCAGGGGA	ACTCGATTCC
10981	TCCCTTTGGA	CCACAGACAT	CCTGGCTTGT	GAATTCACAT	GTCATCTACT	GCTAATCCAT
11041	TGGTAGTATG	TGGCTCACAG	AGACACACTA	CAGTCATGGC	CAATGTCAAG	GTAGGACAGA
11101	TGTGAATCAT	TCCCCCAGTC	CTGCTGTTTT	CATGACTAAC	CCTCCTCAGC	ACAGTGACCA
11161	TGAACCTACT	TTTCCCCCTC	TTTTATTTTT	AGAATTGCTG	GAATTTTCTA	TTTTGAGAAA
11221	TAATAGCCTT	GGGCAGCATT	AAACAAAATC	ATCTAGAAAG	CTGGTTTAAA	ATACAGATGG
11281	TTGAGTCAGT	GAAAGAGTGA	GGAATGTCAT	TATTGGCCCC	TCACAGAGGC	TGGCTCACTC
11341	CAGCAGAGGT	GGTTGAAGCT	CTTGGACACG	GGTCAGGTGC	ATAGGAAAGG	TNGTCTGGGA
11401	CACTGAGAAC	CACAATTGAA	CAACAGAAC	TGTTGGCTTT	TTTTTTTTTA	AATGAGTTCT
11461	CAAAAAATGA	CTGGCTAGCT	TAGGCCAATA	CTTCGAGCCA	ACCCAACAGA	ACATTCTTCC
11521	ATTGATTTAT	TCTGGATCTT	CTTTCTAGAC	AATACTGAAC	TGACCCCTTG	TTGGCAGTCT
11581	CAAGTTTGAC	AACATAGGGC	TTTGAACCTG	GCACAAGGTC	CATCACTGTC	ACCCAAGCAT
11641	CCTGGGTGAC	CTTTGGGTTG	GAATATCTTG	GCTAACCTTA	GATATTTTCT	TTGGAGTATC
11701	TTTAGAACAT	CCAGGAAATA	GGGCTTGATT	CTCATCCTGG	GACCACAATA	TAAGTCACCC
11761	TAGAATCCCA	GGAGATCGTG	CAGAGAAACA	AGGATCTCTC	TCGTGTGCA	CCTTCTTCAA

Figure 23 (continued):

```

11821 AGCAGTGAGT AGTGA CTAAACTGAG TTCCCATCTG AGAGTCCACA GGAGGCTTTG
11881 GGGCAAGAAG CAGAGGGAAG GCACTGTTTG TGTGGTAA GTTTTGACTC TAACAAATTT
11941 GAAGACATAG ATGACATTGT CTCAGACTAA CAACAACCTA GACTCATGTG GGTTCTGTTT
12001 AGGGATCAGA TTTTATTCAT CAATGACTTG TCTTAGTGTA TAGAGAAAGG CTTCTACTG
12061 GAGTGTAGGC TCAATAATGA CAGAAGAGAT AGCTATTTCC CCTAGGGACT GTGCTGCTCC
12121 AAGTTTGGTG GAGAAAGGCA GTGGGGAACC TAGATGTGCT CTCTGGGGAG GGGGTCTGAA
12181 GCTGGCTTCA TAGAAGGTGT GAAGTTTTCG TGAAACATCT AAACAGAATT ATAGCTTAGG
12241 AAAGTGAGCA GGCAAGGCAG GGAATGTGTT GCATATGTAT ATGTACATGA ATATATTATG
12301 TTATAGATAC ACACACATTT GAACCTCATT TGCAGATGAC AGAAAATAGG TTATTTTGCC
12361 TCTCTTAAC TCTAAGCACA ATGACTTCCA GTTCCATCCA TTTCCTGAAA TGCCACAATT
12421 TCATTTTTCA TTGTGGCTGA ATAAAATTCC ATTGCAGACT GGGCCCTACT TCATCCACTC
12481 CTGAGGGCAG GCATATCCCC TGGCTCCATT TCTTACCTAT TGTGAAGAGA AGTGAACATG
12541 TCTTGTGAA AGGCAAGCGT GAGAGAGGCA GGCCTAATT GTGGGTTTTT GTTCTTCTT
12601 CCTGCTATGA CTCTCCATT GTGAGAACCA AAGATCGATA AAAGCCGCCA CCATGAAAGC
12661 CATCTTAATC CCATTTTTAT CTCTCTGAT TCCGTTAACC CCGCAATCTG CATTCGCTCA
12721 GAGTGAGCCG GAGCTGAAGC TGGAAAGTGT GGTGATTGTC AGTCGTCATG GTGTGCGTGC
12781 TCCAACCAAG GCCACGCAAC TGATGCAGGA TGTCACCCCA GACGCATGGC CAACCTGGCC
12841 GGTAAGACTG GGTGGCTGA CACCGCGCGG TGGTGAGCTA ATCGCCTATC TCGGACATTA
12901 CCAACGCCAG CGTCTGGTAG CCGACGGATT GCTGGCGAAA AAGGGCTGCC CGCAGTCTGG
12961 TCAGGTCGCG ATTATTGCTG ATGTCGACGA GCGTACCCGT AAAACAGGCG AAGCCTTCGC
13021 CGCCGGGCTG GCACCTGACT GTGCAATAAC CGTACATACC CAGGCAGATA CGTCCAGTCC
13081 CGATCCGTTA TTTAATCCTC TAAAACTGG CGTTTGCCAA CTGGATAACG CGAACGTGAC
13141 TGACGCGATC CTCAGCAGGG CAGGAGGGTC AATTGCTGAC TTTACCGGCG ATCGGCAAAC
13201 GGC GTTTCGC GAACTGGAAC GGGTGCTTAA TTTTCCGCAA TCAAAC TTGT GCCTTAAACG
13261 TGAGAAACAG GACGAAAGCT GTTCATTAAC GCAGGCATTA CCATCGGAAC TCAAGGTGAG
13321 CGCCGACAAT GTCTCATTA CCGGTGCGGT AAGCCTCGCA TCAATGCTGA CGGAGATATT
13381 TCTCTGCAA CAAGCACAGG GAATGCCGGA GCCGGGGTGG GGAAGGATCA CCGATT CACA
13441 CCAGTGGAAC ACCTTGCTAA GTTTGCATAA CGCGCAATTT TATTTGCTAC AACGCACGCC
13501 AGAGGTTGCC CGCAGCCGCG CCACCCCGTT ATTAGATTTG ATCAAGACAG CGTTGACGCC
13561 CCATCCACCG CAAAAACAGG CGTATGGTGT GACATTACCC ACTTCAGTGC TGTATTATCG
13621 CGGACACGAT ACTAATCTGG CAAATCTCGG CGGCGCACTG GAGCTCAACT GAGCGTCTCC
13681 CCGTCAGCCG GATAACACGC CGCCAGGTGG TGAACGCTG TTTGAACGCT GGCCTCGGCT
13741 AAGCGATAAC AGCCAGTGGA TTCAGGTTTC GCTGGTCTTC CAGACTTTAC AGCAGATGCG
13801 TGATAAAACG CCGCTGTCAT TAAATACGCC GCCCGGAGAG GTGAAACTGA CCCTGGCAGG
13861 ATGTGAAGAG CGAAATGCGC AGGGCATGTG TTCGTTGGCA GGTTTTACGC AAATCGTGAA
13921 TGAAGCACGC ATACCCGCTT GCAGTTTGTA AGGTACCCGG GGATCACAAC TTGCCCTCTG
13981 AAGAGGAAGA ACAGAAGGAT GCCACAATC TCCTGCTGGC TACTCTCCAG TGGTTTCATC
14041 TTACTTCTGA TGGCATTTC CTCTAGAAAG TGCTACTATC ATCCACACAT TTCTACCTGA
14101 GACCACCCAA AGGACCCTCC CAAATCTCT TCTCTCTGA GTAGTCTCCA CACCTGTTAC
14161 CACCATCCCA GAATTAATA CCTAACTGCA CTCTGGCGTG TGACTTGCTT CAGTCTTGC
14221 AATAAGAGTT GTTGGCAGTG CCAGGCGTGG TGGCGCACGC CTTTAATTC AGCACTTGGG
14281 AGGCAGAGGC AGGCGGATTT CTGAGTTCTG GGCACGCTG GTCTACAGAG TGAGTTCCAG
14341 GACAGCCAGG GCTATACAGA GAAACCTGTG GTCGAAAAAC CAAAAAAGTT
14401 GTTGGCAGAG TGTGGGTTAT ATACCAGGTG GAGATTTCAA ATGAGTGGCT GAAGCTGTAG
14461 CCAGAAGGAA CTTAGAGGAT AGCTCATAAC TTAAGAGAA ATGTAGAGAG TAGCAGAAAC
14521 ATTGAGAGAG TGGGCACACA GCCACTGTGT GAATGTGGCA GAACACAATC CAGCCAGCTA
14581 TACATGCATA AGTGTATATT GGCGCCATCC TGACTGATGA GACACAGGAA AACAGATAGA
14641 CGGGGTTAGG TGGCCATGGC CTTTCTGCTG TGCCTCTTCC TAAGGGTCAT CTCAAGACCT
14701 TATGCTCTCT TAACCTTTCC ATTGCTACTT AGCTTCTAGA TATCACCTCC AGATTAGTCT
14761 CCTTGGGTAC ATCAGTGATC CTGGTGATAT CCAGGGCTTC CTGATTCCAT CTTTGTCTA
14821 GAGGCTGCAA CTAAAGAGGT TTTCTTAATA CTTACACACC TGATGCCAAA AGGAAGACAC
14881 AGAAGTTCAC AGAGGTGAAG GTATTCTAGT AGGACATACA GTGAGCAAGC ATCAGGGTCC
14941 GGATTATCTG ACTCTACTCT AACTTTTATG TAAATGTGCT TTATGCCATT AACACTGTCA
15001 TTCCTGTGCT TCAGCTCTGG GAGACTCCCA AGCACTCTTA GGCACAAGCC ACAATTAAGG
15061 GACTCTGACA CTCTGCATTG ATTAATTAGC ATGGTGGTCT CTATGTTTCC AGATTCATGA
15121 TTGTTTCACT TTCCATATAG GCTATGAAGG GTGTGAGGAA ATTTTGTGG GACAGAATTG
15181 GAGGCAATCC ACCTCTCTCA GGAAGCCTCT ATCTGGAATA GCTTACAACT CAGGGACAGT
15241 AACTGTAGGC CCAGTCCTTG GTGTCCAAAA TGGGTTTTAT GGTTTGAATC TGCAAAGCCT

```

Figure 23 (continued):

```

15301 TCCATGTGCT CAAAGGTTTG AACATGGAGC CTCCTCCTGG TAACACTGTA TTGGAGGCTT
15361 TTGAGACTGG ATGCTCTTTG GTCCCATGTT TTGCTACATC ATCTGTCAAG ATATGACCCA
15421 GGCATGCTAC CAGCTACCAC AGACTATGCC TCTCCAGCTT TCATGTTCTC CCCACCATGA
15481 TAGACTTGTA TCTCCTAAAA ATGGAATCAA AGCAAACCTT TCCTGCATTA AGTTTTTTTT
15541 TTTCTGTTAA GTGTTTGGTC ACAGGGACAA GAAAACACTC AATACAGATA ATTAGTACCA
15601 GAGTTGAGGT TCATTGCTCT AGCAAGTTGG ATCAAATTTT TAGGGCTTTG GAACTGATTT
15661 ATAAGAGACA TGTAAGAGAG TCTGAAGCTG TGGGCTACAG AAGTGTCAAC AGTTTTTAAG
15721 AATAGTTTAA TACACCATGG GAATTGTGAA AATCAGAATG CTCACACAAA GGCAGACAGG
15781 AAAACGTGAG CATGTGGCGT GTGAGAGGGC ATAAGAAGGA ACCTAGGGGG AAATGAGCTA
15841 GAAGCCATTC GGCTACGTTA GGAACGTGT GTGGCTGTGC TTGGCCCATG CCCTGGCAAT
15901 CTGAATGAGG CCAAATTTTA AAGGAGTGGA CTAACCTGAT TGTCAGAGAA AATATCAAGA
15961 CAGACCACCA CTCAGGCTAT GCCGTGTTTG TGACCGACCA GCTACTCTTA GCCAGCTCTA
16021 TTGTGAAATT CCAGAGCAAT TATCAGAGCA TGAAGATACA TACAGTTTAG TGAAGTAAGG
16081 GGTGTGGGTC CCTAAGTGGA TGGTGATAA ATCTATGTAG GTGATGCCTA AGTGACACTT
16141 GATAATCCAA AATATCAGCA ATGTGGAATG TCTTCCAAGG AGACCTGTAG ACACACATTT
16201 TAGAACTTTG CTCATGGCTG TAATAAATAG CTAGCTAGAA ATCATTTTCT GAAGAGGTTA
16261 GTCTGAGTTA CGGTTCCAGG GCAAACATTC AGTGATGGCA AGGAAGGCAT TGCAGTCAGG
16321 AGCCAAAGGT CAGCTGGTCA CATTGCATCA AGAGTAGAGA GTCAGAGTGT GAGTAGAAAG
16381 AGGATACAGG TTATAAAACC TCACTGTCCA CTCTCAGCAA TCCATTTTCT CCTAAAAGGC
16441 TTTACCTTCT AAAGATTTTA GTCTTCAAAA CCAGTACCAG TAGCCTGGGA ACAAAGTTG
16501 AAACAAATGA GCCTTTGTGG GGCATTTTAC ACTTAAACA GGGCATCACC TAGGAGGAGC
16561 CCTGTGTGCA GTAGGAAGTG TGGCCTCTGT GTCAGGAATG CTCAGGCTAA TAAGGGGTCC
16621 TCTATCTGAG GGACCCTATG AAGATTCAAC AAGTAGTTGT GAGAATTCCC TGTAAATGGA
16681 TGCTACCAAT TTGACATTTG TAGACCTGCT ATGTGTGCT TCTTTATTGG GCTCTCCCAT
16741 CTCCCAACTT TCCAACCCAT ATTCCACATT AATCCCTTCC ACCACCATGC AACACTAGGT
16801 AGGAGAGAAG GAAGGTTAGA AGAGAAAGTG GGTATAGATC TATTTAGACT ACTTCCTGCT
16861 GATTAGGGGC AAGTCCAATC GTCATTGTCA GGATACCTCC AACCAGCAAC CAGCAAACCA
16921 GCAAATCAGA AACAGCAAAA GCAGCCAACA AGGCAGCACT AACCAGCAGG ATTGGGGTCG
16981 GTAGCGTGGG AGCAGTCACT ACTGGTCTTC TCATGGCTTT GGCATTAATA CTCTCTCAAG
17041 AAATCCGTA ATTTTTTCCC CACCACCTGA AATCCGTAA TTTTAAATGC AAATATCTA
17101 CAGCTGGCAA AAATCACATC TCTCCTAGAG CACAAGACAA ATCATAGTTA CTGGCTATTT
17161 GCAATCTGAA GCATCTCAAT ATCCCACACC TGGGATTAAA ACAAAAACAT ATTCACATCA
17221 CATACTGTT TTTTTTTTCC AATTTTTTAT TAGGTATTTT CTTTATTTAC ATTTCAAATG
17281 CTATCCCGAA AGTCCCCTAT ACCCTCCAC CTCCCTGCTC CCCTACACAC CCCTCCAC
17341 TTTTGTACCC TGGAGTTCCC CGGTACTGGG GCATATAAAG TTTGCAAGAC CAAGGGGCCT
17401 CTCTTCCAG TGATGGCCGA CTAAGCCATC TTCTGCTACA TATGCAGATA GAGACACGAG
17461 CTCTGGGGGT ACTAGTTAGT TCATATTGTT GTTCCACCTA TAGGGTCGCA GACCCCTTCA
17521 GCTCCTTGGG TACTTTGTCT AGCTCCTCCA CTGGGGGCTC TGTGTTTTAT CTAATAGATG
17581 ACTGTGAGCA TCCACTTCTG TATTTGACAG GCACTGGCCT AGCGTCACAT GAGCCAGCTA
17641 TATCAGGGTC CTTTCAGCAA AACCTTGCTG GCATGTGCAA TAGTGTCTGC GTTTGGTGGT
17701 TGATTATGGG ATGGATCCAC TAGTTCTAGA GC

```

//